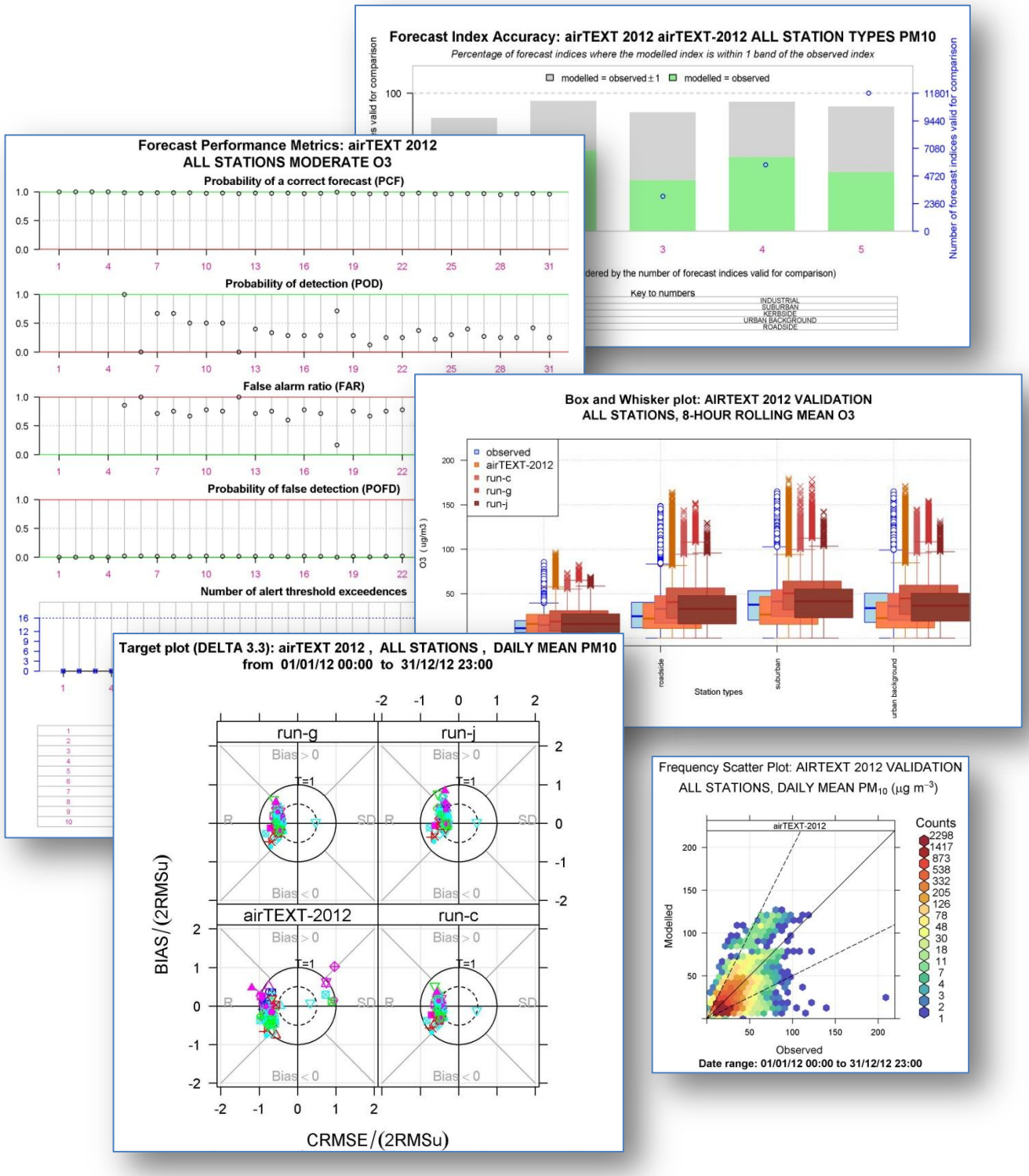
Myair Toolkit for Model Evaluation User Guide

Version 3.0, June 2013



Contents

1 Introduction

Regional and municipal governments are increasingly interested in providing services to assess and forecast local- and city-level air quality. Air quality forecasts on these scales can be disseminated to health services and the public in terms of air quality alerts, to inform and warn 'at-risk' groups about impending pollution episodes and provide advice. Local air quality modelling is critical in assessment of air quality against the EC air quality directive as it can provide high resolution maps of concentration where the population is most dense and allows the investigation of proposed mitigation measures on short or long time scales. Understanding the benefits, limitations and performance of individual models, the input data required of them as well the extent of the options available to them is often lacking. Setting standard evaluation criteria and comparing model capabilities in a structured way is therefore a crucial task.

This toolkit has been developed under the local forecast model evaluation support work package of the EU's 7th Framework, PASODOBLE project. It draws on existing best practice such as the EU Joint Research Council's (JRC) FAIRMODE initiative on model evaluation [1] and the openair project tools [2].

It is a simple-to-install, user-friendly environment that guides the user through the process of evaluating model predictions of local air quality and investigating the model performance. It runs on Windows, Mac or linux operating systems.

The toolkit can take modelled data from regional or local scale models as input. Observed data are *in situ* time series data. Missing data are handled if they are indicated by a standard value. As output, the toolkit creates plots of the model performance in predicting concentrations and predicting alerts with respect to defined thresholds, for single or multiple sites, single or multiple pollutants and single or multiple modelled data sets. Results can be classified by the type of monitoring site and the pollutant for each modelled data set. The diagnosis of model performance for individual sites and individual pollutants produces time series plots, scatter plots and analyses with respect to month, day of the week and hour of the day. All the plotted data are also exported to data files to provide an audit trail and make the data available for further analysis and visualisation.

2 Getting started

The Toolkit can be used on Windows, Mac or linux operating systems and does not require any software to be purchased. Before using the Toolkit you will need to install some free software: R and some R packages, java and RGGRunner; this will just take a few minutes. Detailed installation instructions are given in sections 2.1-2.4.

2.1 Install R

The MyAir Toolkit for Model Evaluation version 3.0 is compatible with R version 2.15.3.

Follow these step-by-step instructions to download and install R from the internet:

1. Go to http://www.r-project.org/
2. Select CRAN from the links on the left-hand side of the page
3. Choose a CRAN mirror for your locality (in the UK, choose the mirror for the University of Bristol) and click on the link
4. Under 'Download and install R' click on the link for your operating system
5. Click on 'base'
6. Under 'Other builds' click on 'Previous releases'
7. Click on 'R 2.15.3 (March 2013)'
8. Click on 'Download R...' to download the install program
9. Run the install program, taking care to install R in a directory where you have write privileges.

NOTE: If you do not have direct access to the internet from your computer, for example you access the internet through a university network, then when you install R, instead of accepting all the defaults, at the Setup screen choose not to accept all the defaults and when offered, choose 'Internet2' as the internet option. This will force R to use the same proxy settings used by Internet Explorer. The defaults for all other options can be accepted.

2.2 Install R packages

Follow these instructions to install the required packages for R:

1. Open R
2. From the Packages menu, select 'Install package(s)...'
3. Pick a CRAN mirror, preferably one that is geographically nearest to you, as this is likely to be the fastest. In the UK, choose the Bristol mirror.
4. Scroll down to find 'openair' from the list of packages, click on it and click 'OK' to install openair
5. Repeat steps 2 and 4 for the following packages (you won't be asked to choose a mirror again): ncdf, GEOmap, fields, latticeExtra, maps, plyr, akima, hexbin, lattice and nlme

2.3 Install Java

Follow these instructions to download and install Java (needed for RGGRunner):

1. Go to http://www.java.com/
2. Follow the instructions to download and install Java (version 5 or higher)

2.4 Install RGGRunner

For Windows and Linux operating systems:

1. In the 'Installs' subdirectory, open rggrunner_win_linux.zip
2. Extract all the files to a suitable directory (e.g. C:\Program Files on Windows; the extraction will put all files in a subdirectory called 'rggrunner')
3. Make a desktop shortcut to rggrunner.exe in the 'bin' subdirectory; this is the RGGRunner program

For Mac operating systems:

1. Find rggrunner_macosx.zip in the 'Installs' subdirectory
2. Extract all files to a suitable directory
3. After extraction, double click to mount the image
4. After mounting, double click again on the "RGG" icon to run the application

3 Using the toolkit

The Myair Toolkit for Model Evaluation consists of four tools:

1. Screening Questionnaire

   This tool gives structured advice on your proposed evaluation.

2. Data Input

   This tool processes your modelled and observed concentration data, saving the processed data in an R workspace and (optionally) a CSV file.

3. Model Evaluation

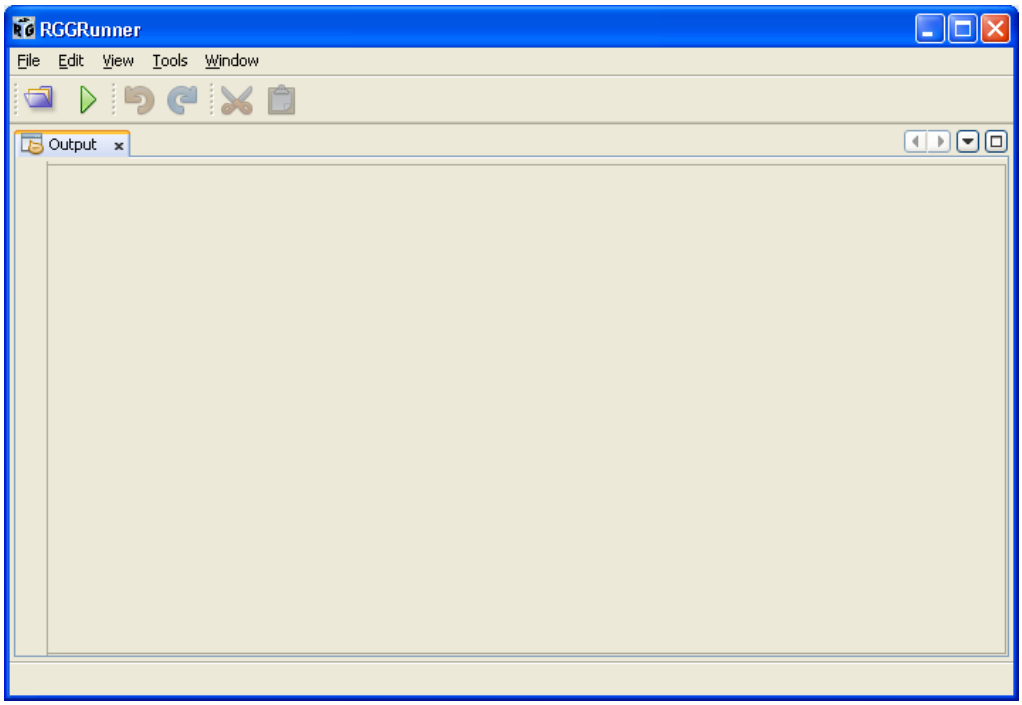
   This tool takes in the workspace saved by the Data Input tool and evaluates all or some of the data, producing graphs and (optionally) CSV files.

4. Model Diagnostics

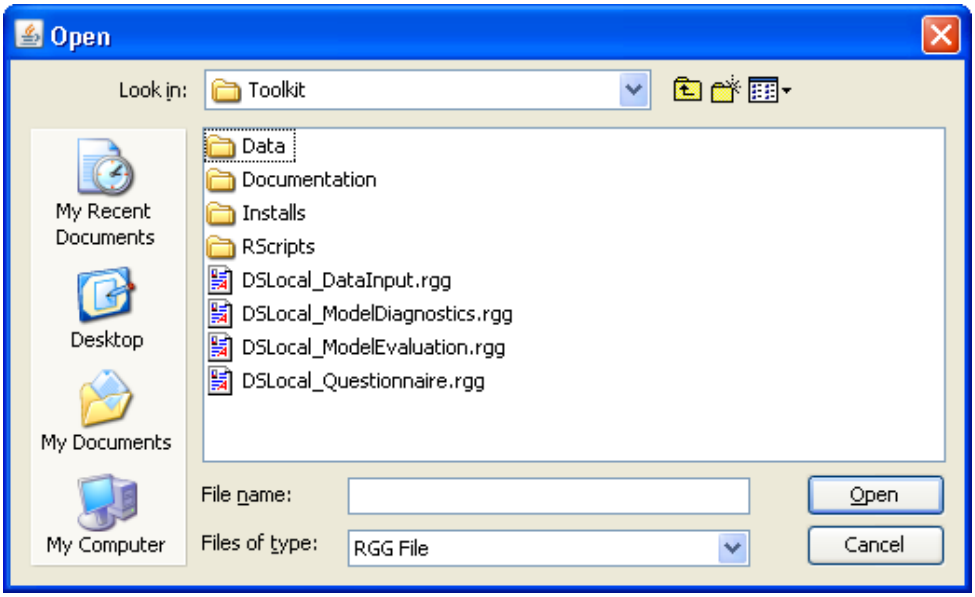
   This tool takes in the workspace saved by the Data Input tool and produces diagnostic graphs for one station and one pollutant at a time.

This section of the User Guide gives step-by-step instructions for using each tool in the Toolkit.

Step 1: Start RGGRunner



Step 2: **Click on File, Open RGG and browse to select each of the four Toolkit tools in turn**



Step 3: **You have now loaded the Myair Toolkit for Model Evaluation**

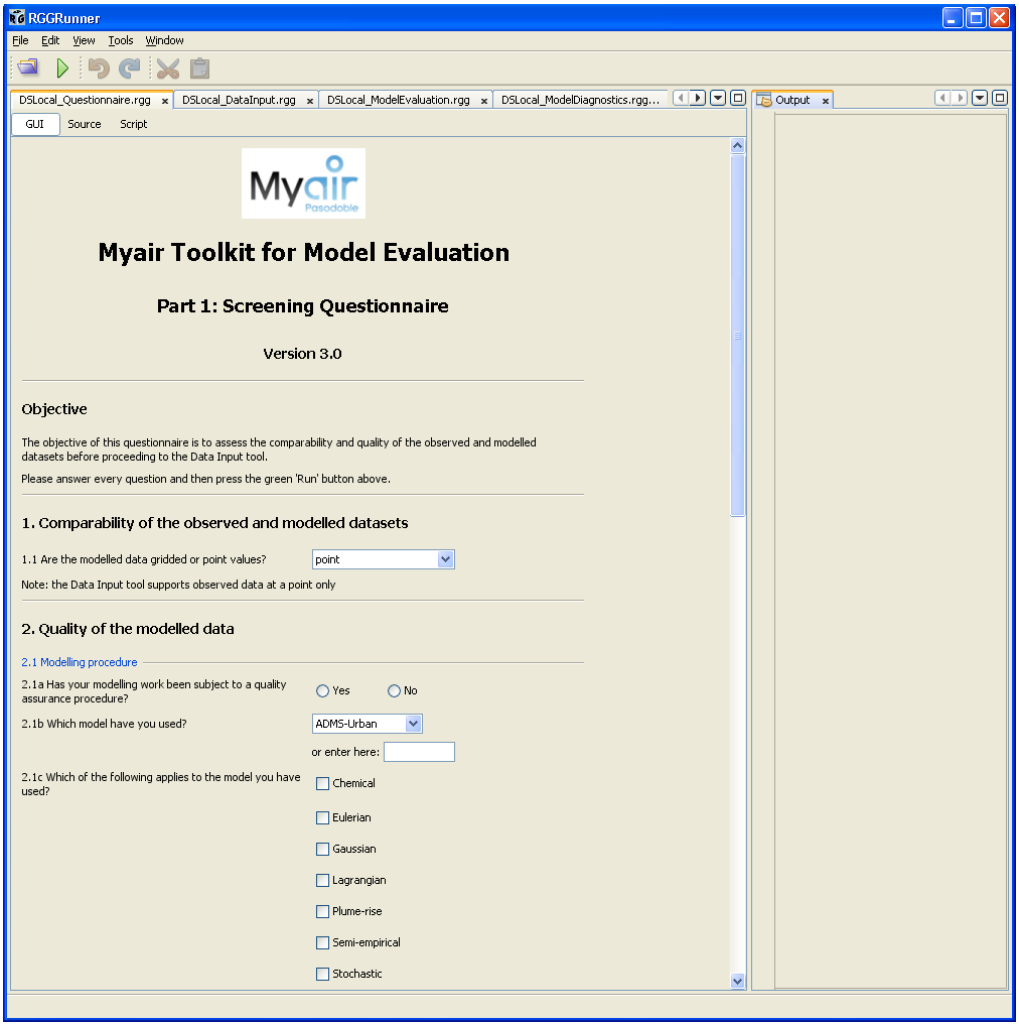


Figure 3.1 The Myair Toolkit for Model Evaluation

3.1 Questionnaire tool

The questionnaire is a screening tool which asks you questions about the data you want to evaluate and offers structured advice based on your answers.

Answer each question in turn, then click the Play button  .

3.2 Data Input tool

The data input tool processes your modelled and observed data, saving it in an R workspace file, which later can be imported into the Model Evaluation and Model Diagnostics tools.

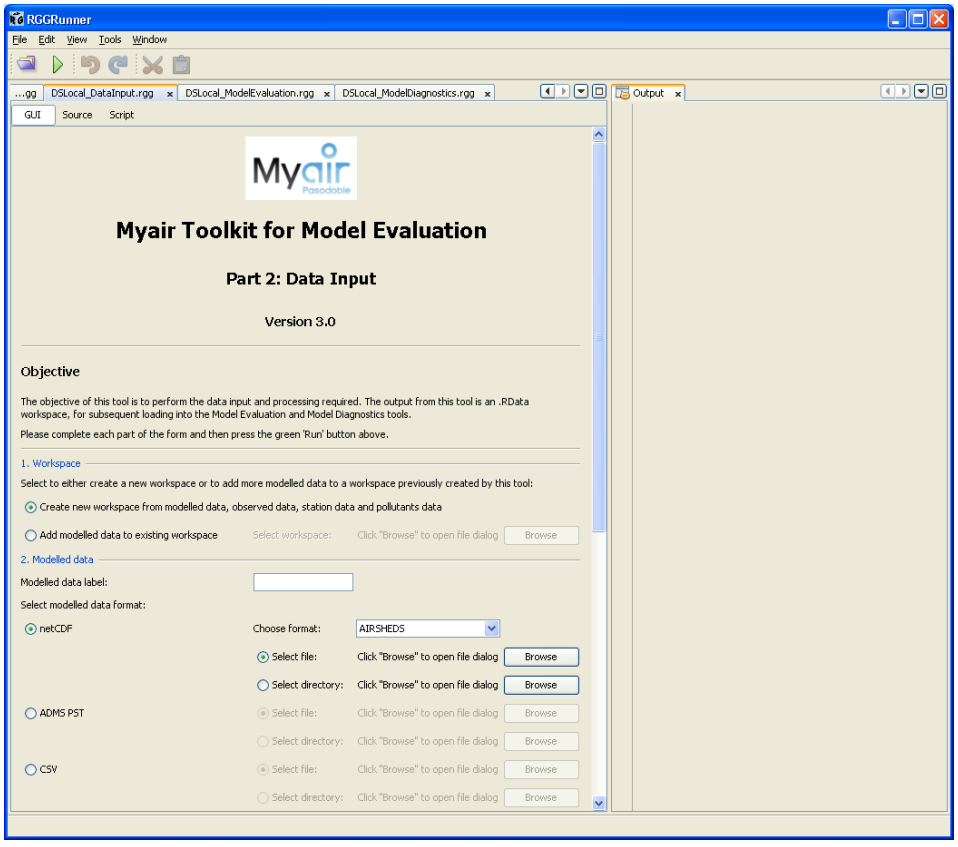
The tool supports both gridded and point modelled data. Gridded data are interpolated to the monitoring station locations.

In situ observed data can either be automatically downloaded from the internet and imported (UK only) or input using simple-format CSV files.

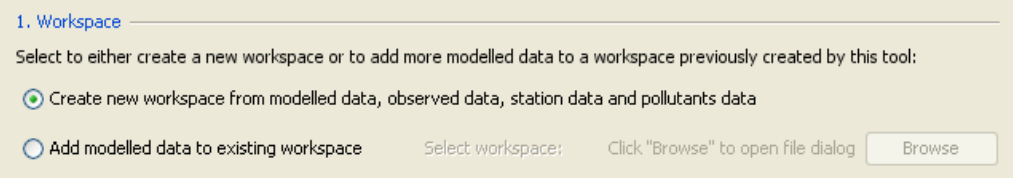
The Toolkit supports the evaluation of multiple modelled data sets, either from multiple models or from multiple runs of the same model, over the same time period. Initially, a new R workspace is created with one modelled data set and an associated observed data set. Further modelled data sets can then be added to the saved R workspace. There is no limit to the number of models that the Data Input tool will load into a saved R workspace, but memory issues may be encountered if the number of models is large.

Tip: In the Myair Toolkit for Model Evaluation installation directory you will find a 'DataSamples' sub-directory. Here you will find sample files in the formats recognised by the Toolkit, and a ReadMe.txt file describing each file.

Step 1: In RGGRunner, select the Data Input tool



Step 2: **Select whether you wish to create a new workspace or add modelled data**



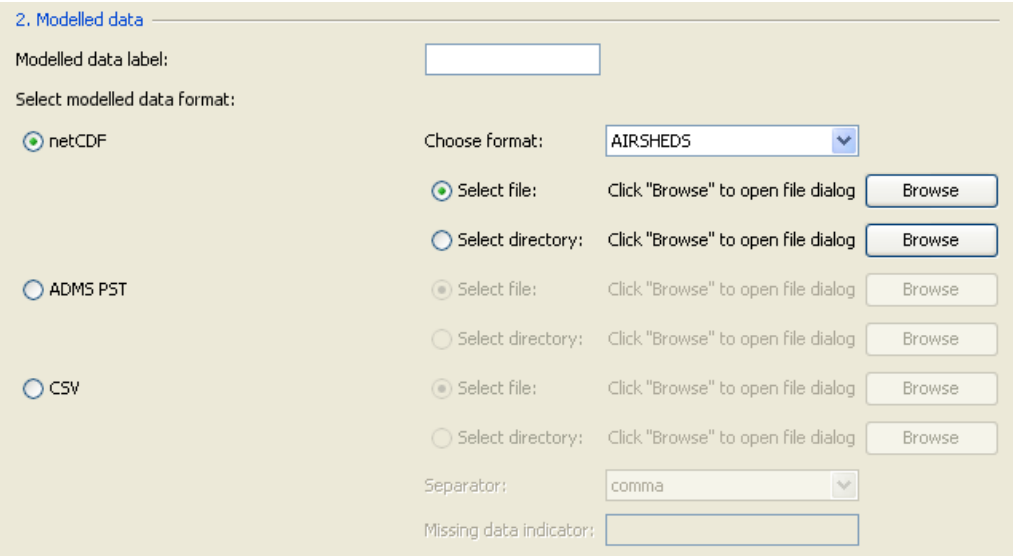
You can select to create a new R workspace or to add a modelled data set to an existing R workspace.

Note that, when creating a new workspace, the data range of the modelled data set should cover the start and end time you wish to evaluate.

When adding a modelled data set to an existing R workspace, the R workspace must have been created by version 3.0 of the Data Input tool.

When adding multiple modelled data sets to a workspace, the modelled data averaging time and statistic must be the same for all modelled data sets. Refer to Table 3.1 for details.

Step 3: **Select your modelled data.**



You need to label the modelled data set being read in to the tool.

This label will be used in the subsequent plots to identify the modelled data set. The label should be alpha-numeric and not include any special characters such as '&', '!' or commas.

You can select either a single file or a whole directory of data files.

The Data Input tool supports the following formats for modelled data:

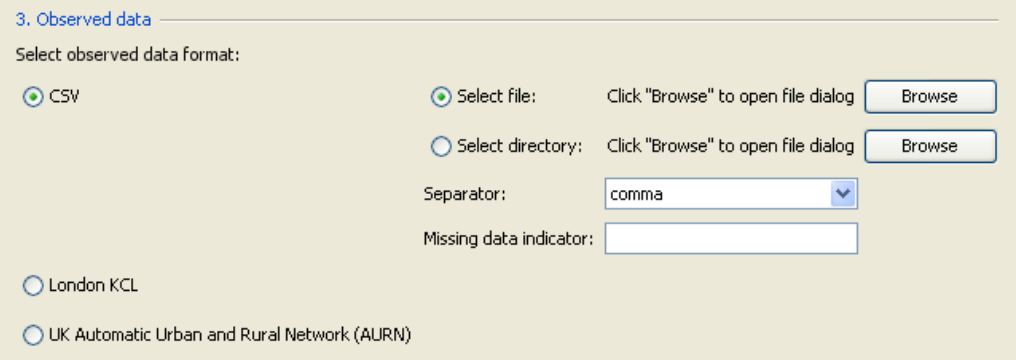
- netCDF files
    - AIRSHEDS: gridded modelled data from the PASODOBLE IC-AIRSHEDS work package

      This option supports output from the PASODOBLE Web Coverage Service (WCS) and output from individual IC-AIRSHEDS partners

- MACC Ensemble: gridded data from the MACC regional ensemble air quality product
- CMAQ: gridded modelled data from the CMAQ model

- ADMS PST: point receptor output format from the ADMS suite of atmospheric dispersion models
- CSV: standard text file
  - Refer to section 5.1 for details of the required format
  - Select the separator used in your CSV files – either comma or semicolon
  - Enter the missing data indicator used in your CSV files (e.g. -999 or NA)

Step 4: **If creating a new workspace, select your required observed data option.**



The Data Input tool supports the following formats for observed data:

- CSV: standard text file
  - Refer to section 5.1 for details of the required format
  - Select the separator used in your CSV files – either comma or semicolon
  - Enter the missing data indicator used in your CSV files (e.g. -999 or NA)
  - Select either a single CSV file or a whole directory of CSV files
- London KCL

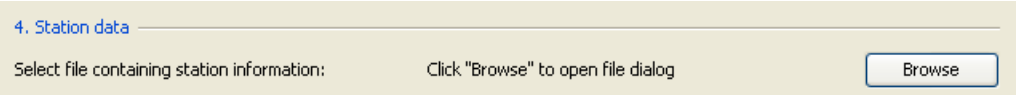
  This option requires internet access

  - King's College London (KCL) maintain a network of 123 monitoring sites in Greater London
  - Data for all required stations and pollutants will be downloaded and imported for a time period to match the modelled data
- UK Automatic and Rural Network (AURN)

  This option requires internet access

  - The UK AURN is the national network of automatic monitoring stations around the UK
  - Data for all required stations and pollutants will be downloaded and imported for a time period to match the modelled data

Step 5: **If creating a new workspace, select station data file**



Your station data file should be a CSV file with a comma separator, containing a list of all the stations for which you wish to process data. Each station should have a station type, and, if you have selected one of the gridded netCDF modelled data formats, you also need to include the station longitude and latitude.

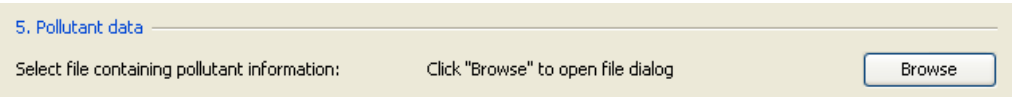
	A	B	C	D	E
1	name	type	latitude	longitude	
2	BG1	Suburban	51.56375	0.177928	
3	BG2	Suburban	51.52939	0.132889	
4	BG3	Roadside	51.54043	0.074451	
5	BL0	Urban Bac	51.52206	-0.12578	
6	BN1	Roadside	51.61467	-0.17658	
7	BN2	Urban Bac	51.59189	-0.20596	
8	BQ5	Industrial	51.47769	0.190833	
9	BQ6	Industrial	51.47849	0.185298	
10	BQ7	Urban Bac	51.49476	0.134384	
11	BQ8	Urban Bac	51.49476	0.134384	
12	BT1	Suburban	51.58962	-0.27549	
13	BT4	Roadside	51.55247	-0.25806	
14	BT5	Industrial	51.55265	-0.24875	
15	BT6	Roadside	51.53779	-0.24776	

stations_KCL_latlong

Figure 3.2 Example of a station data file

If you are using either the London KCL or UK AURN observed data options then the station names must match the station codes used by those networks.

Step 6: **If creating a new workspace, select pollutant data file**



Your pollutant data file should be a CSV file with a comma separator, containing a list of all the pollutants for which you wish to process data. For each pollutant a set of parameters must be set. **It is very important these parameters are set correctly for your data.** Refer to Table 3.1 for details of the information required.

	A	B	C	D	E	F	G	H	I	J	K
1	pollutant	output.units	conv.ugm3.ppb	obs.alias	obs.units	obs.avg.time.hours	obs.statistic	mod.alias	mod.units	mod.avg.time.hours	mod.statistic
2	no2	ug/m3	0.52	no2	ug/m3	1	mean	no2_conc	ug/m3	1	mean
3	o3	ug/m3	0.5	o3	ug/m3	1	mean	o3_conc	ug/m3	1	mean
4	pm10	ug/m3	1	pm10	ug/m3	1	mean	pm10_conc	ug/m3	1	mean
5											
6											

pollutants_macc

Figure 3.3 Example of a pollutant data file

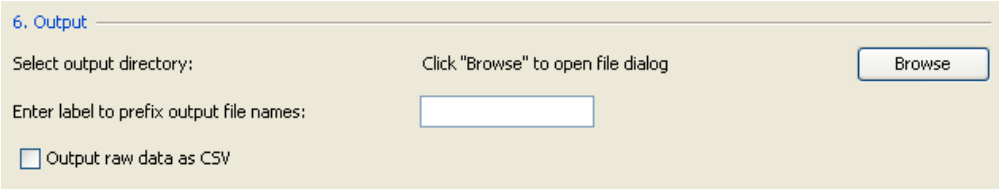
Notes:

- The 'obs.alias' and 'mod.alias' settings will be the same as 'pollutant' unless the pollutant name within the observations or modelled data are different to 'pollutant', in which case use the aliases to make sure the correct fields are extracted from the data.
- The observed data averaging time and statistic must either be the same as for the modelled data, or 1 hour mean (then it will be recalculated to the same averaging time and statistic as the modelled data).
- When adding multiple modelled data sets to a workspace, the modelled data averaging time and statistic must be the same for all modelled data sets.

Column header	Description	Allowed values
pollutant	Name of the pollutant to be used in all output	*n/a*
output.units	Concentration units to be used in all output	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
conv.ugm3.ppb	Conversion factor from µg/m^3 to ppb (used for unit conversions)	Any numeric value
obs.alias	Pollutant name as it appears in the observed data	*n/a*
obs.units	Units that apply to the observed data	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
obs.avg.time.hours	Averaging time in hours of the observed data (minimum 1 hour)	An integer value, minimum 1
obs.statistic	Statistic that applies to the observed data	'max', 'mean' or 'rolling mean'
mod.alias	Pollutant name as it appears in the modelled data	*n/a*
mod.units	Units that apply to the modelled data	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
mod.avg.time.hours	Averaging time in hours of the observed data (minimum 1 hour)	An integer value, minimum 1
mod.statistic	Statistic that applies to the observed data	'max', 'mean' or 'rolling mean'

Table 3.1 Details of the pollutant data CSV file columns. n/a indicates that values are unrestricted.

Step 7: **Select output directory**



Browse to select the directory where you want the workspace to be saved.

Step 8: **Enter a label for output files**

The workspace will be saved with this name (with extension *.RData*), and all output files subsequently produced by the Model Evaluation and Model Diagnostics tools using this workspace will also be labelled with this name.

Check the box if you wish the whole time series of modelled and observed data for every station and pollutant to be output to a CSV file.

Warning: this file may be large.

Step 9: **Click the green Play button** 

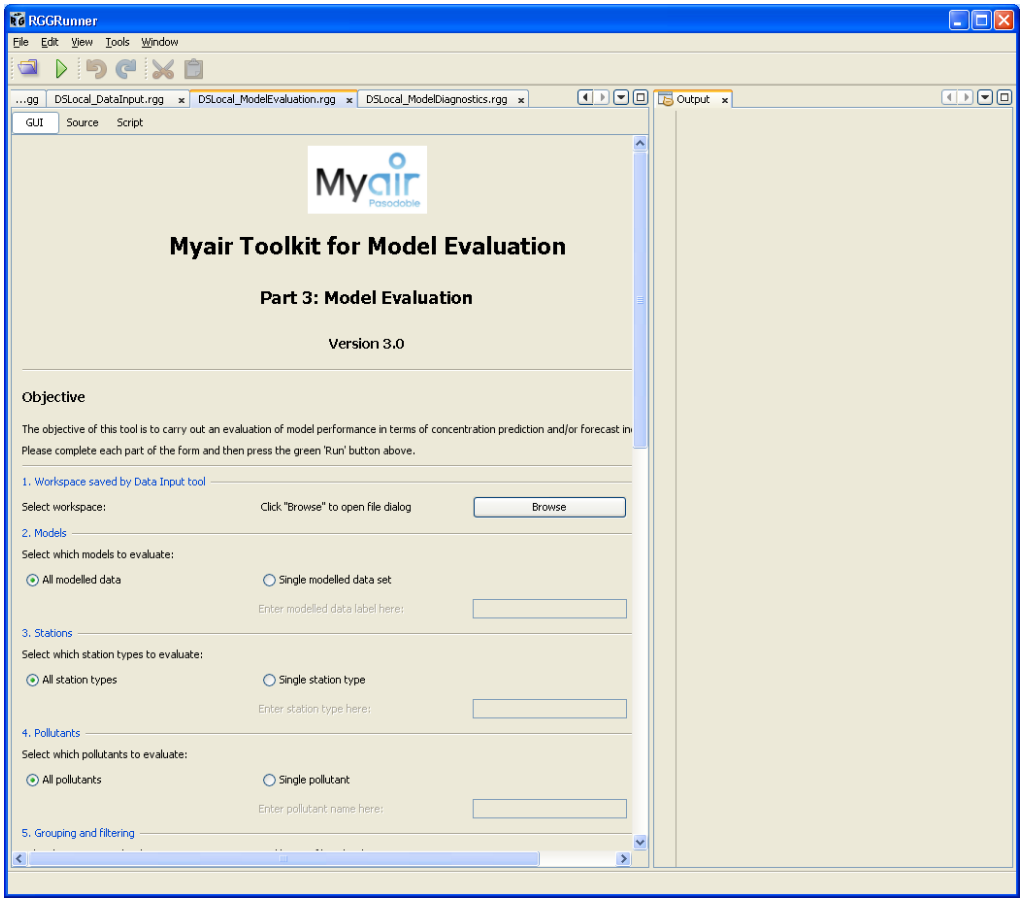
Refer to section 4.1 for details about the output from the Data Input tool.

3.3 Model Evaluation tool

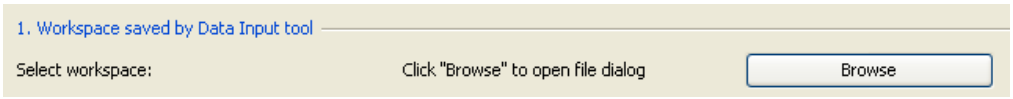
This section gives a step-by-step guide to using the Model Evaluation tool. For details of each graph type, please refer to section 4.2.

Tip: In the Myair Toolkit for Model Evaluation installation directory you will find a 'DataSamples' sub-directory. Here you will find sample files in the formats recognised by the Toolkit, and a ReadMe.txt file describing each file.

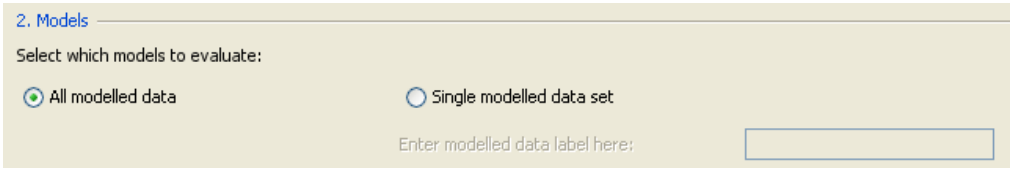
Step 1: **In RGGRunner, select the Model Evaluation tool**



Step 2: **Browse to select the workspace previously saved by the Data Input tool**



Step 3: **Select which modelled data sets to evaluate**



You can either evaluate all modelled data sets or just one modelled data set. If you only want to evaluate one modelled data set, enter its label (as specified in the Data Input tool) in the box provided.

Tip: If you don't know the labels of the modelled data sets that are available, type any text into the box. When you run the tool it will declare that no data are available for that modelled data set and give you a list of the available labels.

Step 4: Select which station types to evaluate

You can either evaluate all station types or just stations of one type. If you only want to evaluate stations of one type, enter the name of this station type in the box provided.

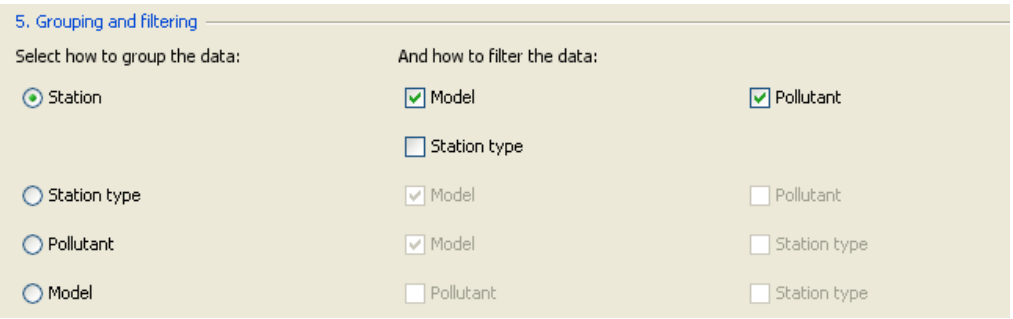
Tip: If you don't know which station types are available, type any text into the box. When you run the tool it will declare that no data are available for that station type and give you a list of the available station types.

Step 5: Select which pollutants to evaluate

You can either evaluate all pollutants or just one pollutant. If you only want to evaluate one pollutant, enter the name of this pollutant in the box provided.

Tip: If you don't know which pollutants are available, type any text into the box. When you run the tool it will declare that no data are available for that pollutant and give you a list of the available pollutants.

Step 6: Select how to group and filter the data



'Grouping' describes what each data point on a plot represents. The Model Evaluation tool supports grouping the data by:

- Station[1]
- Station type
- Pollutant
- Model (or modelled data set)

[1] It is noted that, for the frequency scatter plot, conventional scatter plot and QQ plot options, the option to group by station should be chosen. In these specific cases, the data are plotted for all stations together on the same graph. If data for an individual station is required, refer to section 3.4 for the Model Diagnostics tool.

'Filtering' describes how the data is to be split across plots, either as separate panelled plots on one page or as additional pages (for large plots).

The default option is to group by station and filter by pollutant and model. Data must always be either grouped by or filtered by model (or modelled data set) because the data sets must always be evaluated separately against the observed data.

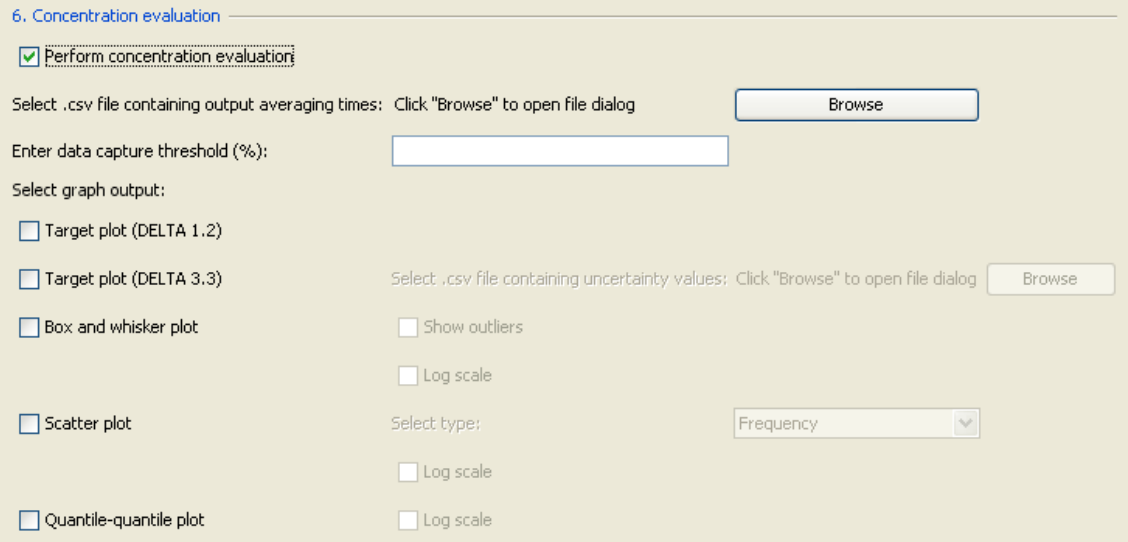
Note that there are limitations on the grouping and filtering variables available for each plot option, appropriate to what each plot presents. For example, the forecast index evaluation plots present normalised variables and so all options are available with the only constraint being that the data must be either grouped or filtered by model. Whereas for concentration evaluation plots, the data must be grouped or filtered by both pollutant and model. Table 3.2presents the available paired options for grouping and filtering for each plot. Refer to sections 4.2.1 and 4.2.2 for details about these graphs.

Plot	Allowed Group	Allowed Filter
Target plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
	Station type	Model and Pollutant
	Pollutant	Model
	Pollutant	Model and Station type
	Model	Pollutant
	Model	Pollutant and Station type
Box and whisker plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
	Station type	Model and Pollutant
	Model	Pollutant and Station type
Scatter plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
Quantile-quantile plot	Station	Model and Pollutant
	Station	Model and Pollutant and Station type
Forecast index plots	Station	Model
	Station	Model and Pollutant
	Station	Model and Station type
	Station	Model and Pollutant and Station type
	Station type	Model
	Station type	Model and Pollutant
	Pollutant	Model
	Pollutant	Model and Station type
	Model	-
	Model	Pollutant
	Model	Station type
	Model	Pollutant and Station type

Table 3.2 Grouping and filtering options for the Model Evaluation tool

Step 7: Concentration Evaluation

This part of the tool compares the modelled concentrations to observed concentrations.



- Check the box to perform the concentration evaluation.
- Select your output averaging times file.

  Your output averaging times file should be a comma-separated CSV file containing one row per pollutant. For each pollutant, the required averaging time and statistic must be set. Table 3.3 below describes these parameters.

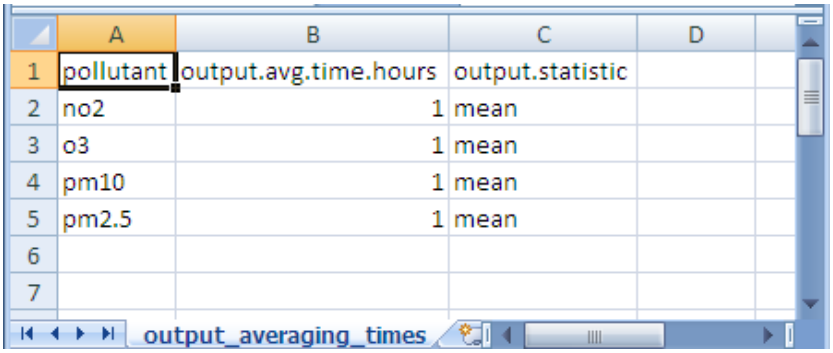


Figure 3.4 Example of an output averaging times file

Column header	Description	Allowed values
pollutant	Name of the pollutant as defined in the pollutant data file used in the Data Input tool	*n/a*
output.avg.time.hours	Averaging time in hours to be applied to the concentration evaluation output (minimum 1 hour)	An integer value, minimum 1
output.statistic	Statistic that applies to the concentration evaluation output	'max', 'mean' or 'rolling mean'

Table 3.3 Details of the output averaging times CSV file columns. n/a indicates that values are unrestricted.

- Enter a data capture threshold to apply to the output averaging process. For example, for an 8-hour rolling mean, a 75% data capture threshold means that at least 6 hours of data must be valid in each period for the averaged data for this period to be considered valid and used in the evaluation.

- Five graph options are available: a target plot from version 1.2 of the FAIRMODE DELTA Tool, a target plot from version 3.3 of the FAIRMODE DELTA Tool, a box and whisker plot, a scatter plot as either a frequency (binned) or conventional scatter plot and a quantile-quantile plot. Check the boxes of the graphs you require. Refer to section 4.2.1 for details about these graphs.

- If you have selected the target plot from DELTA v3.3, select your uncertainties file.

  Your uncertainties file should be a comma-separated CSV file containing one row per pollutant. For each pollutant, the coefficients of the measurement uncertainty given by the FAIRMODE DELTA v3.3 methodology [3] must be specified. Table 3.4 describes these parameters.

- If you have selected the box plot option, check the 'Show outliers' option if you want outliers to be plotted on the graph.

- If you have selected the scatter plot option, select which type of plot you require; a frequency scatter plot or conventional scatter plot.

  Tip: The frequency scatter plot is better suited to larger sets of data, the conventional scatter plot to smaller sets.

- If you have selected any of the box plot, scatter plot or quantile-quantile plot options, check the 'Log scale' option if you wish the numerical axes of the plot to be scaled to $\log_{10}$.

	A	B	C	D	E	F	G	H
1	pollutant	k	ur	LV	alpha	target.units	target.avg.time.hours	target.statistic
2	NO2	2	0.12	200	0.02	ug/m3	1	mean
3	O3	1.4	0.09	120	0.62	ug/m3	8	rolling mean
4	PM10	2	0.138	50	0.027	ug/m3	24	mean
5								
6								

uncertainties

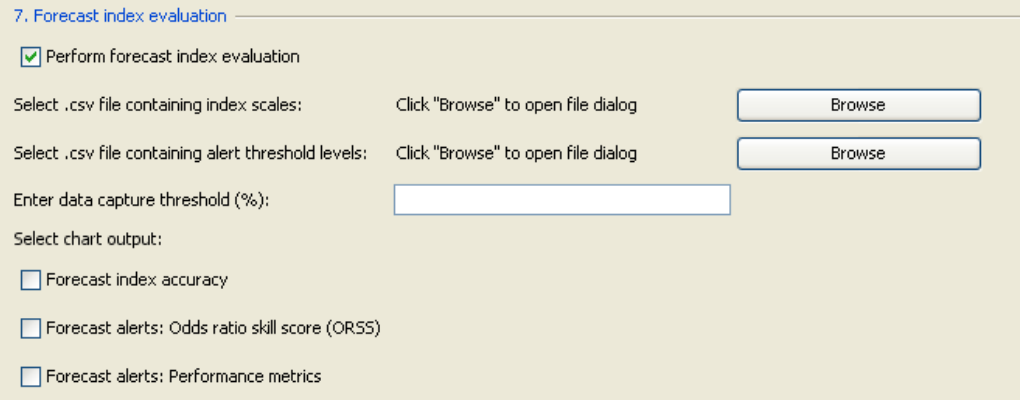
Figure 3.5 Example of an uncertainties file

Column header	Description	Allowed values
pollutant	Name of the pollutant as defined in the pollutant data file used in the Data Input tool	*n/a*
k	Coverage factor, *k*.	Refer to the DELTA v3.3 methodology [3]
ur	Measurement uncertainty, u_r^{LV}	Refer to the DELTA v3.3 methodology [3]
LV	Limit value (or reference value)	Refer to the DELTA v3.3 methodology [3]
alpha	Proportion of the measurement uncertainty that is independent of the limit value.	Refer to the DELTA v3.3 methodology [3]
target.units	Units that apply to the measurement uncertainty	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
target.avg.time.hours	Averaging time in hours that applies to the measurement uncertainty, which should be equal to that applied to the concentration evaluation output (minimum 1 hour)	An integer value, minimum 1
target.statistic	Statistic that applies to the measurement uncertainty, which should be equal to the concentration evaluation output	'max', 'mean' or 'rolling mean'

Table 3.4 Details of the uncertainties CSV file columns. n/a indicates that values are unrestricted.

Step 8: Forecast Index Evaluation

This part of the tool converts both the observed and modelled concentrations to 'forecast indices' and performs an evaluation based on these forecast indices.



- Check the box to perform the forecast index evaluation.
- Select your index scales file.

  Your index scales file should be a comma-separated CSV file containing one row per pollutant. For each pollutant, index-related parameters must be set. Table 3.5 describes these parameters. A minimum of 1 index threshold concentration must be defined.

- Select your alert thresholds file.

  Your alert thresholds file should be a comma-separated CSV file containing one row per alert. Each alert should have a name and an alert threshold. Table 3.6 describes these parameters.

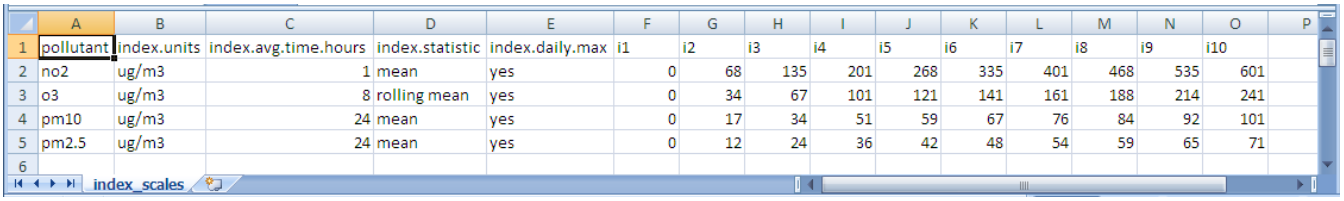
- Enter a data capture threshold to apply to the forecast index averaging process. For example, for an 8-hour rolling mean, a 75% data capture threshold means that at least 6 hours of data must be valid in each period for the averaged data to be valid and used in the evaluation. The threshold will also be applied over each day if the daily maximum is set to “yes” in your index scales file.

- Select the graphs you require

  There are three graph options (for details about these graphs, refer to section 4.2.2):

  1. Forecast index accuracy
  2. Forecast alerts: Odds ratio skill score (ORSS)
  3. Forecast alerts: Performance metrics

  Check the boxes for the graph output you require.



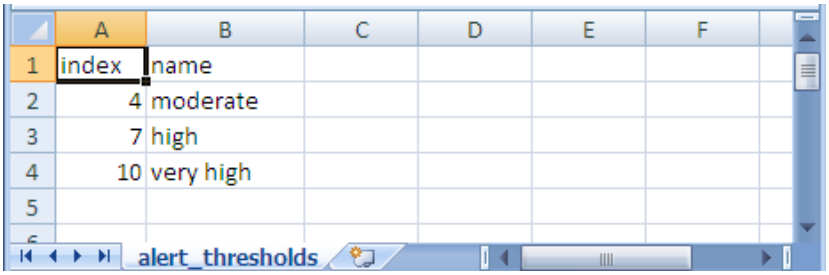
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	pollutant	index.units	index.avg.time.hours	index.statistic	index.daily.max	i1	i2	i3	i4	i5	i6	i7	i8	i9	i10	
2	no2	ug/m3	1	mean	yes	0	68	135	201	268	335	401	468	535	601	
3	o3	ug/m3	8	rolling mean	yes	0	34	67	101	121	141	161	188	214	241	
4	pm10	ug/m3	24	mean	yes	0	17	34	51	59	67	76	84	92	101	
5	pm2.5	ug/m3	24	mean	yes	0	12	24	36	42	48	54	59	65	71	
6																

index_scales

Figure 3.6 Example of an index scales file

Column header	Description	Allowed values
pollutant	Name of the pollutant as defined in the pollutant data file used in the Data Input tool	*n/a*
index.units	Units in which the index threshold concentrations are given	'mol/mol', 'ppb', 'ppm', 'ng/m3', 'ug/m3', 'mg/m3', 'g/m3' or 'kg/m3'
index.avg.time.hours	Averaging time in hours that applies to the index threshold concentrations (minimum 1 hour)	An integer value, minimum 1
index.statistic	Statistic that applies to the index threshold concentrations	'max', 'mean' or 'rolling mean'
index.daily.max	Is the forecast given as the daily maximum of the calculated indices?	'yes' or 'no'
i1	Threshold concentration for index level 1	Any numeric value
i2	Threshold concentration for index level 2	Any numeric value
i{n}	Threshold concentration for index level n	Any numeric value

Table 3.5 Details of the index scales CSV file columns. n/a indicates that values are unrestricted.



	A	B	C	D	E	F
1	index	name				
2	4	moderate				
3	7	high				
4	10	very high				
5						

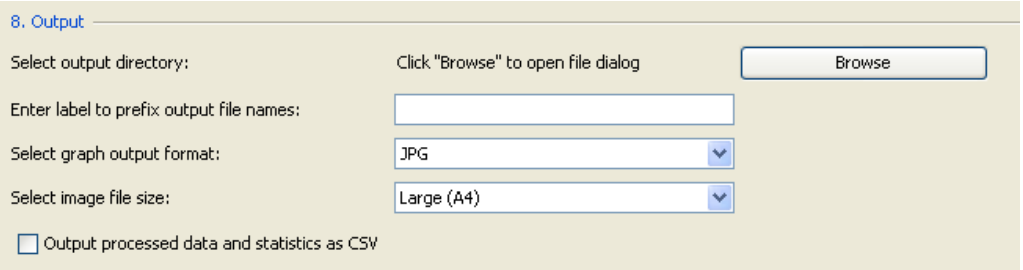
alert_thresholds

Figure 3.7 Example of an alert thresholds file

Column header	Description	Allowed values
index	Threshold index for this alert	Any integer value
name	The name to give to this alert in all output	*n/a*

Table 3.6 Details of the index scales CSV file columns. n/a indicates that values are unrestricted.

Step 9: Select output directory



Browse to select the directory where you want the output files to be saved.

Step 10: Enter label to prefix output file names

The text you enter here will be used to identify your output.

Step 11: Select graph output format

There are three options: JPG, PNG and PDF.

The first two options produce image files that can be imported into other documents. One image file is produced for every graph. The PDF option produces PDF files, with one PDF file per graph type.

Step 12: Select image file size

There are three options: Large (A4), Medium and Small.

These options only apply to the image output formats, JPG or PNG. PDF output is always produced on A4. Reducing the overall image size will increase the proportional size of text. These options provide flexibility to produce graphs for reports or presentations.

Step 13: Select whether to output processed data and statistics as CSV

This is a very useful option that produces CSV files containing all of the numerical data used to create the graphs, also some statistics not shown on the graphs. Refer to section 4.2 for details of the contents of these files.

Step 14: Click the green Play button 

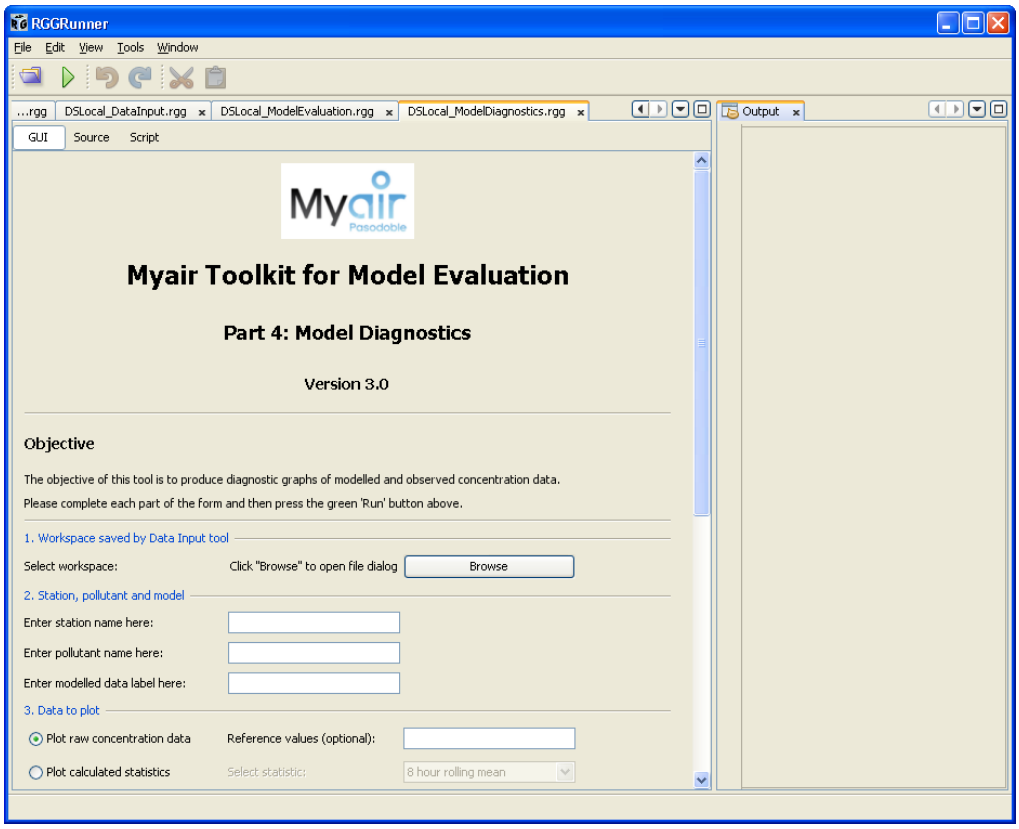
Refer to section 4.2 for details of the output from the Model Evaluation tool.

3.4 Model Diagnostics Tool

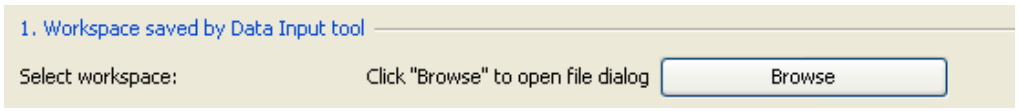
The aim of the model diagnostics tool is to enable further investigation of the performance of the model at one particular monitoring station for one particular pollutant.

Tip: In the Myair Toolkit for Model Evaluation installation directory you will find a 'DataSamples' sub-directory. Here you will find sample files in the formats recognised by the Toolkit, and a ReadMe.txt file describing each file.

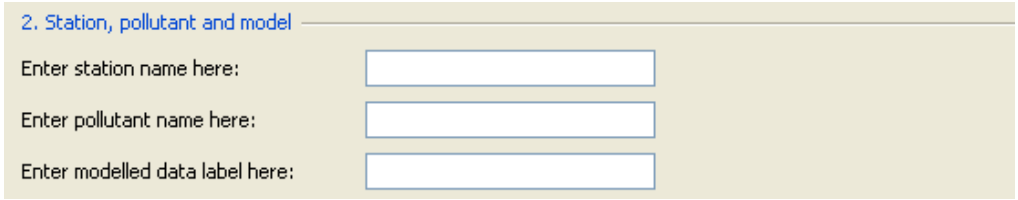
Step 1: In RGGRunner, select the Model Diagnostics tool



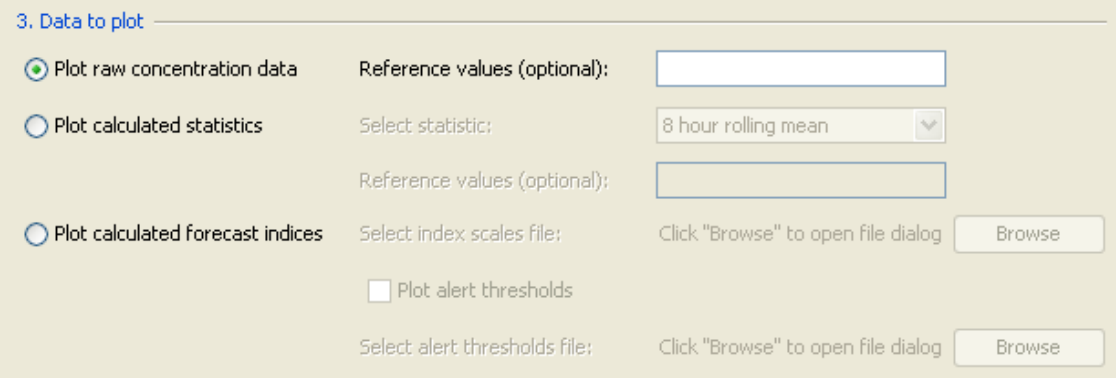
Step 2: Browse to select the workspace previously saved by the Data Input tool



Step 3: Enter station, pollutant and modelled data label



Step 4: **Select which data to plot**



There are three options here:

1. Plot raw concentration data

   This plots the data as it was entered, in the averaging time of the modelled data.

   If required, one to ten reference values may also be plotted on the graph. To enter more than one reference value, separate the values with a comma, e.g. '50, 100'.

2. Plot calculated statistics

   This plots the selected statistic, from a choice of the following: 8 hour rolling mean, 8 hour mean, daily mean and daily maximum.

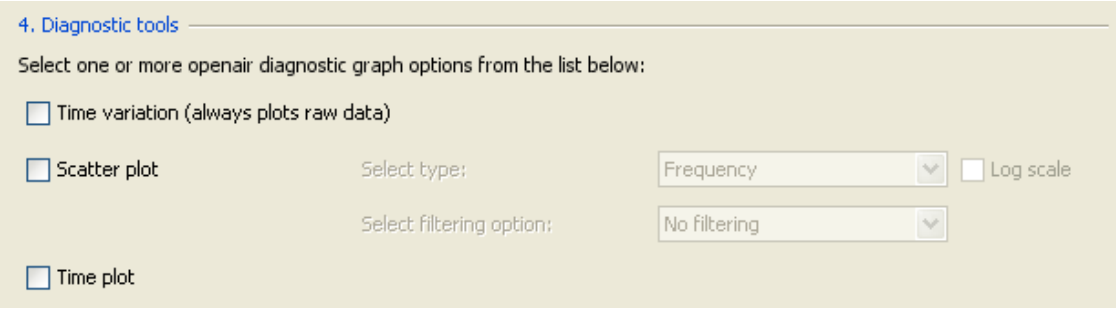
   If required, one to ten reference values may also be plotted on the graphs. To enter more than one reference value, separate the values with a comma, e.g. '50, 100'.

3. Plot calculated forecast indices

   This converts both the modelled and observed data to indices (as defined in the index scales file described in section 3.3, step 8, and detailed in Table 3.5).

   If required, alert thresholds defined in the alert thresholds file (described in section 3.3, step 8) can be plotted on the graphs as reference lines.

Step 5: **Select which openair graph options you require**



There are three options here (for full details on the graphs refer to section 4.3):

1. Time variation

   This plots averages by weekday, month and hour. The settings in the 'Data to plot' section do not apply to this graph; it uses the raw data as it was entered into the Data Input tool.

2. Scatter plot

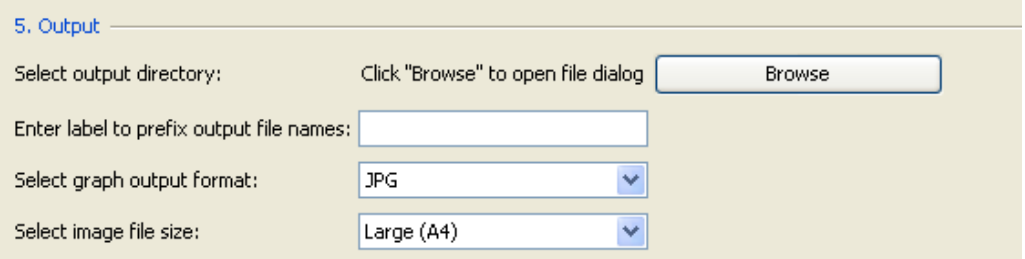
   Select which type of plot you require; a frequency scatter plot shows the frequency of occurrence of each point whereas a conventional scatter plot shows one point per pairwise modelled-observed data point. The scatter plot can be scaled to $\log_{10}$ and filtered by weekday, month or season.

   Tip: The frequency scatter plot is better suited to larger sets of data, the conventional scatter plot to smaller sets.

3. Time plot

   This is a time series plot of both modelled and observed data.

Step 6: Select output directory



Browse to select the directory where you want the output files to be saved.

Step 7: Enter label to prefix output file names

The text you enter here will be used to identify your output.

Step 8: Select graph output format

There are three options: JPG, PNG and PDF.

The first two options produce image files that can be imported into other documents. The PDF option produces PDF files.

Step 9: Select image file size

There are three options: Large (A4), Medium and Small.

These options only apply to the image output formats, JPG or PNG. PDF output is always produced on A4. Reducing the overall image size will increase the proportional size of text. These options provide flexibility to produce graphs for reports or presentations.

Step 10: Click the green Play button 

Refer to section 4.3 for details of the output from the Model Diagnostics tool.

4 Output

4.1 Data Input Tool Output

4.1.1 R Workspace

This is an R workspace containing all the data imported by the Data Input tool, processed and ready for import into the Model Evaluation and Model Diagnostic tools.

This workspace can also be loaded into R to explore the data further.

4.1.2 CSV file

This contains the time series of concentrations (in output units as defined in the pollutants information file) for every station and pollutant for which both monitored and modelled data are available, with the same averaging time and statistic as the modelled data.

4.2 Model Evaluation Tool Output

4.2.1 Concentration Evaluation Output

4.2.1.1 Target plot (DELTA 1.2)

The 'Target' plots produced by the concentration evaluation part of the Model Evaluation tool are similar to the plot produced by the FAIRMODE DELTA tool [1]. These plots can be output as defined by either version 1.2 or version 3.3 of the DELTA tool. This section describes output in line with version 1.2 of the DELTA tool.

For DELTA version 1.2, the metrics calculated by the tool for each monitoring station and pollutant and shown on the target plot are:

Centralised root-mean-square error (CRMSE):

$$\mathrm{CRMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left((M_i - \bar{M}) - (O_i - \bar{O})\right)^2}$$

$\bar{M}$ is the mean modelled concentration, $\bar{O}$ is the mean observed concentration.

Mean bias (BIAS):

$$\mathrm{BIAS} = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$$

Standard deviation of the observations (σ_{obs}):

$$\sigma_{obs} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - \bar{O})^2}$$

Standard deviation of the modelled data (σ_{mod}):

$$\sigma_{mod} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - \bar{M})^2}$$

Target (T):

$$T = \sqrt{\left(\frac{\text{BIAS}}{\sigma_{obs}}\right)^2 + \left(\frac{\text{CRMSE}}{\sigma_{obs}}\right)^2}$$

The target plot shows $\frac{\text{BIAS}}{\sigma_{obs}}$ against $\left[\frac{\text{CRMSE}}{\sigma_{obs}}\right]_{\sigma}$ where

$$\left[\frac{\text{CRMSE}}{\sigma_{obs}}\right]_{\sigma} = \begin{cases} -\frac{\text{CRMSE}}{\sigma_{obs}} & \text{if } \sigma_{mod} > \sigma_{obs} \\ \frac{\text{CRMSE}}{\sigma_{obs}} & \text{if } \sigma_{obs} > \sigma_{mod} \end{cases}$$

The radial distance to a data point on the target plot is equal to T for that station. The smaller the value of T, the better the concentration prediction for that station.

The black circle represents T=1, the dark green circle represents T=0.8, the light green circle represents T=0.65 and the dotted line represents T=0.3. In DELTA v1.2, 0.8 is the 'criteria' value, 0.65 is the 'goal' value and 0.3 is the uncertainty.

Figure 4.1 shows an example of target plot (DELTA 1.2) output from the concentration evaluation part of the model evaluation tool (all pollutants, all station types).

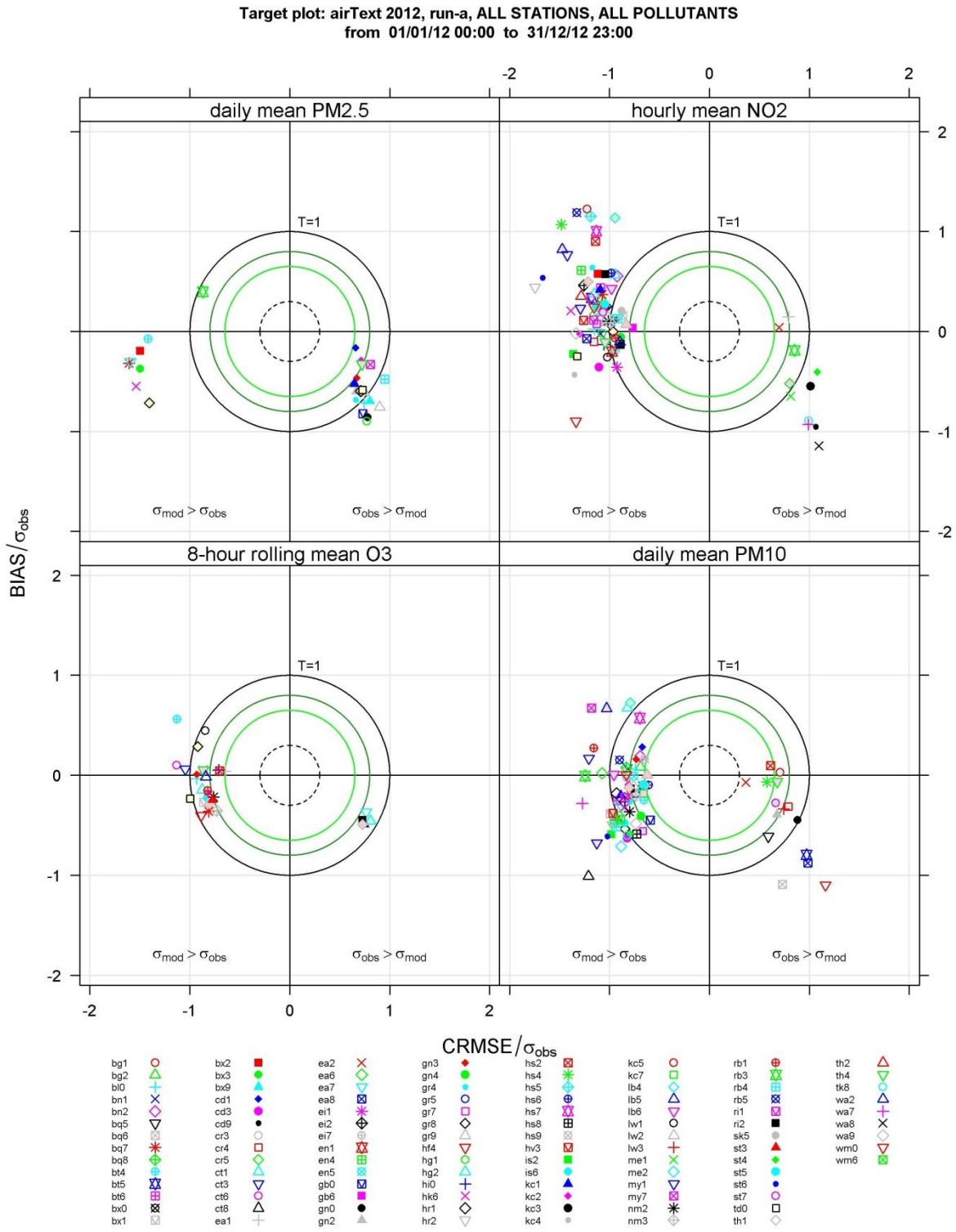


Figure 4.1 Example of a target plot (DELTA 1.2) from the concentration evaluation part of the Model Evaluation tool (grouped by station and filtered by model and pollutant for a single model 'run-a')

4.2.1.2 Target plot (DELTA 3.3)

The 'Target' plots produced by the concentration evaluation part of the Model Evaluation tool are similar to the plot produced by the FAIRMODE DELTA tool [1]. This section describes output in line with version 3.3 of the DELTA tool [3].

For DELTA version 3.3, the metrics calculated by the tool for each monitoring station and pollutant and shown on the target plot are:

Centralised root-mean-square error (CRMSE):

$$\mathrm{CRMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left((M_i - \bar{M}) - (O_i - \bar{O})\right)^2}$$

$\bar{M}$ is the mean modelled concentration, $\bar{O}$ is the mean observed concentration.

Mean bias (BIAS):

$$\mathrm{BIAS} = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$$

Root-mean-square of the observation measurement uncertainty (RMSu):

$$\mathrm{RMS_u} = k u_r^{LV}\sqrt{\left((1-\alpha)(\bar{O}^2 + \sigma_o^2)\right) + \alpha LV^2}$$

where σ_o^2 is the standard deviation of the observed data, *LV* is the limit value of interest and the other coefficients, k, u_r^{LV} and α, are derived for a specific pollutant from measurement data, as described in the FAIRMODE DELTA v3.3 methodology [3] and associated papers [4,5]. An example uncertainties file is provided in the data examples, using the values reported by DELTA v3.3:

	k	u_r^{LV}	α	*LV* (ug/m3)
NO2	2	0.120	0.020	200
O3	1.4	0.090	0.620	120
PM10	2	0.138	0.027	50

Target (*T*):

$$T = \sqrt{\left(\frac{\mathrm{BIAS}}{2\mathrm{RMS_u}}\right)^2 + \left(\frac{\mathrm{CRMSE}}{2\mathrm{RMS_u}}\right)^2}$$

The target plot shows $\frac{\mathrm{BIAS}}{2\mathrm{RMS_u}}$ against $\left[\frac{\mathrm{CRMSE}}{2\mathrm{RMS_u}}\right]_{\mathrm{NMSD}}$ where

$$\left[\frac{\mathrm{CRMSE}}{2\mathrm{RMS_u}}\right]_{\mathrm{NMSD}} = \begin{cases} -\dfrac{\mathrm{CRMSE}}{2\mathrm{RMS_u}} & \text{if } \dfrac{\mathrm{NMSD}}{\sqrt{2(1-R)}} < 1 \\ \dfrac{\mathrm{CRMSE}}{2\mathrm{RMS_u}} & \text{if } \dfrac{\mathrm{NMSD}}{\sqrt{2(1-R)}} > 1 \end{cases}$$

R is the Pearson's correlation coefficient and NMSD is the normalised mean standard deviation $(\sigma_m - \sigma_o)/\sigma_o$.

The radial distance to a data point on the target plot is equal to T for that station. The smaller the value of T, the better the concentration prediction for that station.

The black circle represents T=1, which represents the performance criteria and should be fulfilled by at least 90% of stations. The dashed black circle represents T=0.5. The grey lines separate quadrants representative of the dominant error in the modelled data, from positive or negative BIAS, or from standard deviation (SD) or correlation (R) errors in the CRMSE.

Figure 4.2 and Figure 4.3 show examples of output from the concentration evaluation part of the model evaluation tool with different grouping and filtering options.

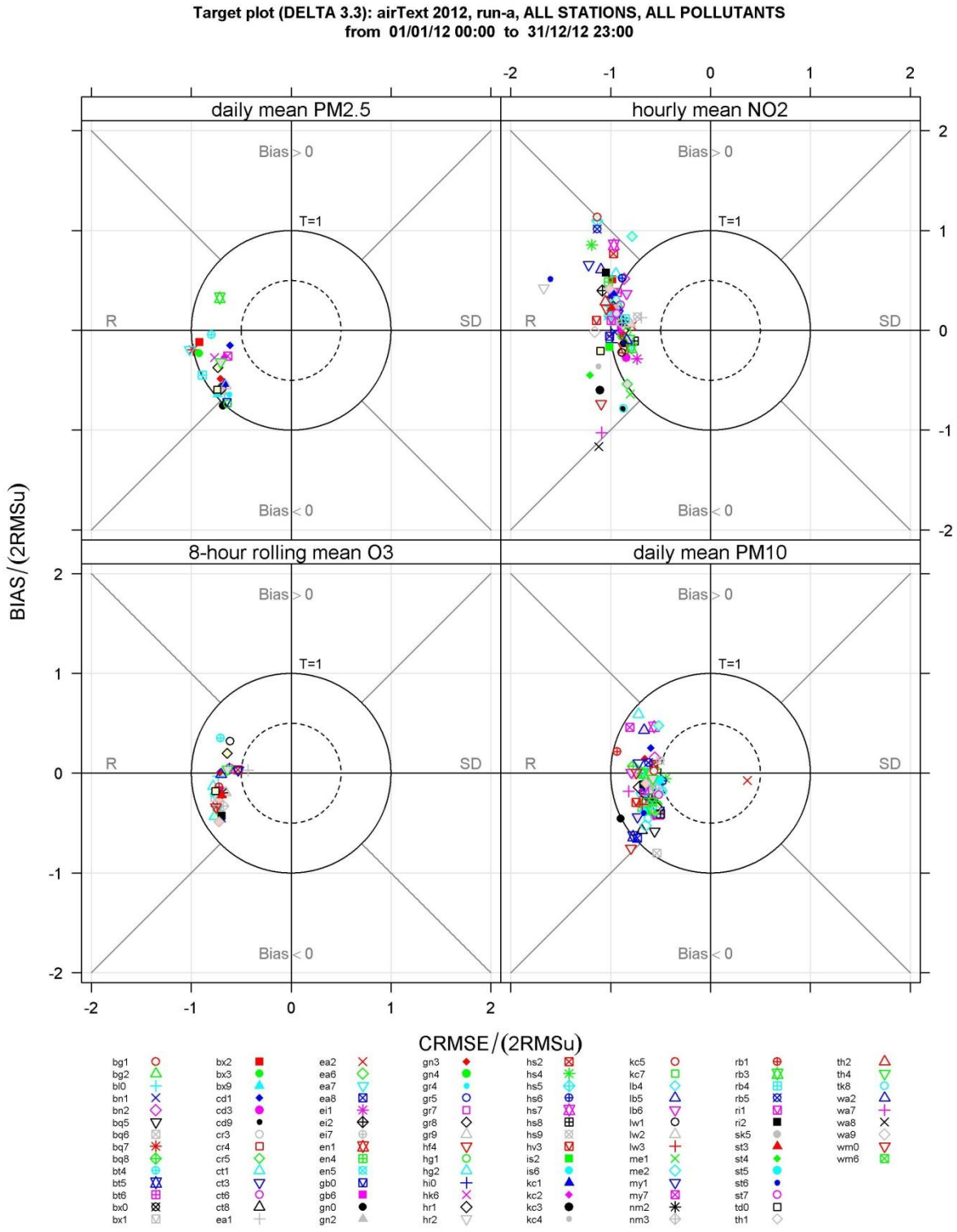


Figure 4.2 Example of a target plot (DELTA 3.3) from the concentration evaluation part of the Model Evaluation tool (grouped by station and filtered by model and pollutant for a single model 'run-a')

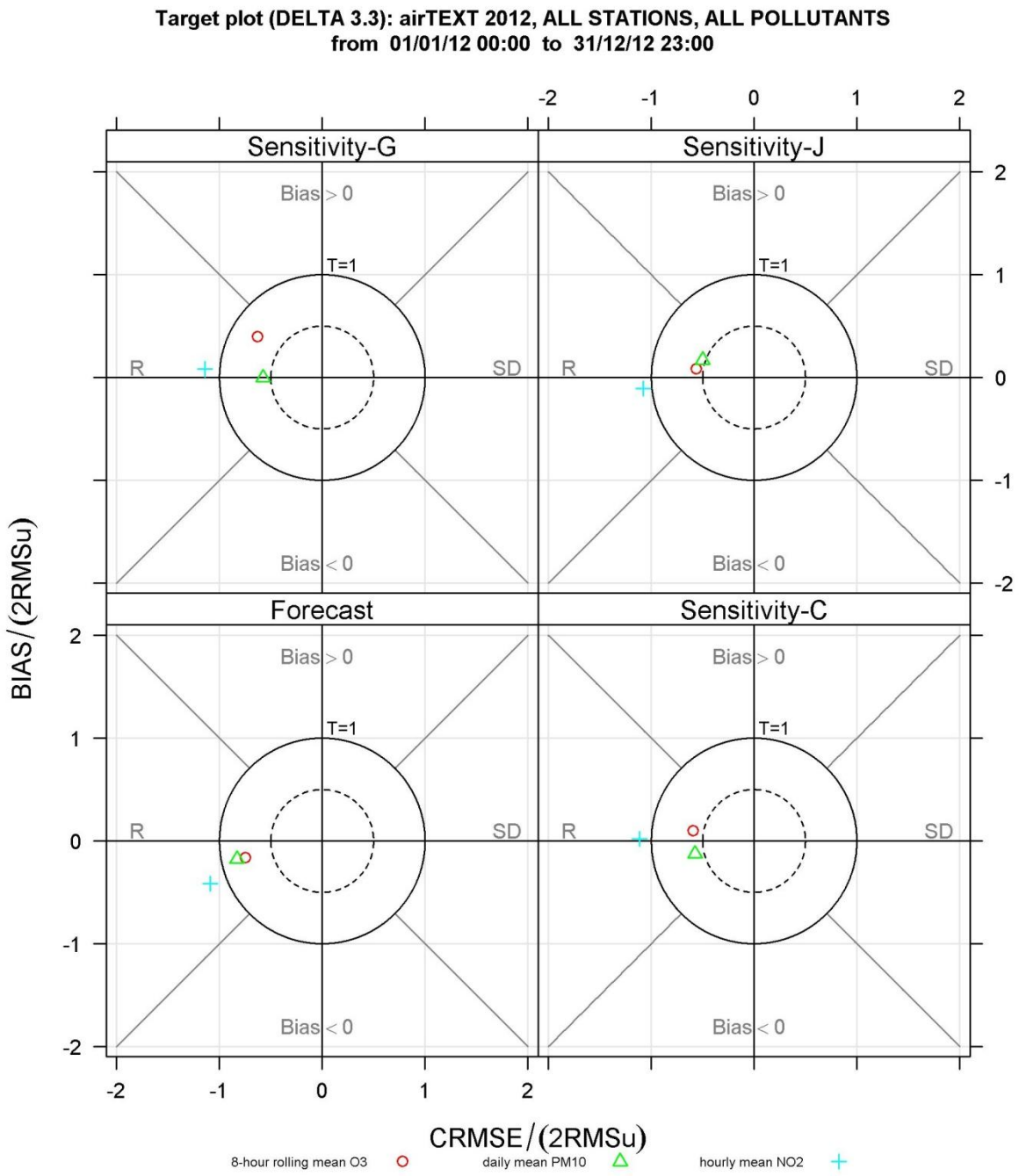


Figure 4.3 Example of a target plot (DELTA 3.3) from the concentration evaluation part of the Model Evaluation tool (grouped by pollutant and filtered by model)

4.2.1.3 Box and Whisker plot

The Box and Whisker plot shows 5 pieces of information for each station, for each of the modelled and observed datasets:

1. The lower whisker
2. The 25th percentile (the lower quartile – the lower end of the box)
3. The 50th percentile (the median – the horizontal line inside the box)
4. The 75th percentile (the upper quartile – the upper end of the box)
5. The upper whisker

The inter-quartile range (IQR) is defined as the 75th percentile minus the 25th percentile, i.e. the length of the box.

The lower whisker is defined as the lowest concentration value still within 1.5xIQR of the lower quartile. The higher whisker is defined as the highest concentration value still within 1.5xIQR of the upper quartile.

Optionally, the outliers lying outside of the upper and lower whisker can also be plotted, and the plot can be displayed on a $\log_{10}$ scale.

An example box and whisker plot is shown in Figure 4.4.

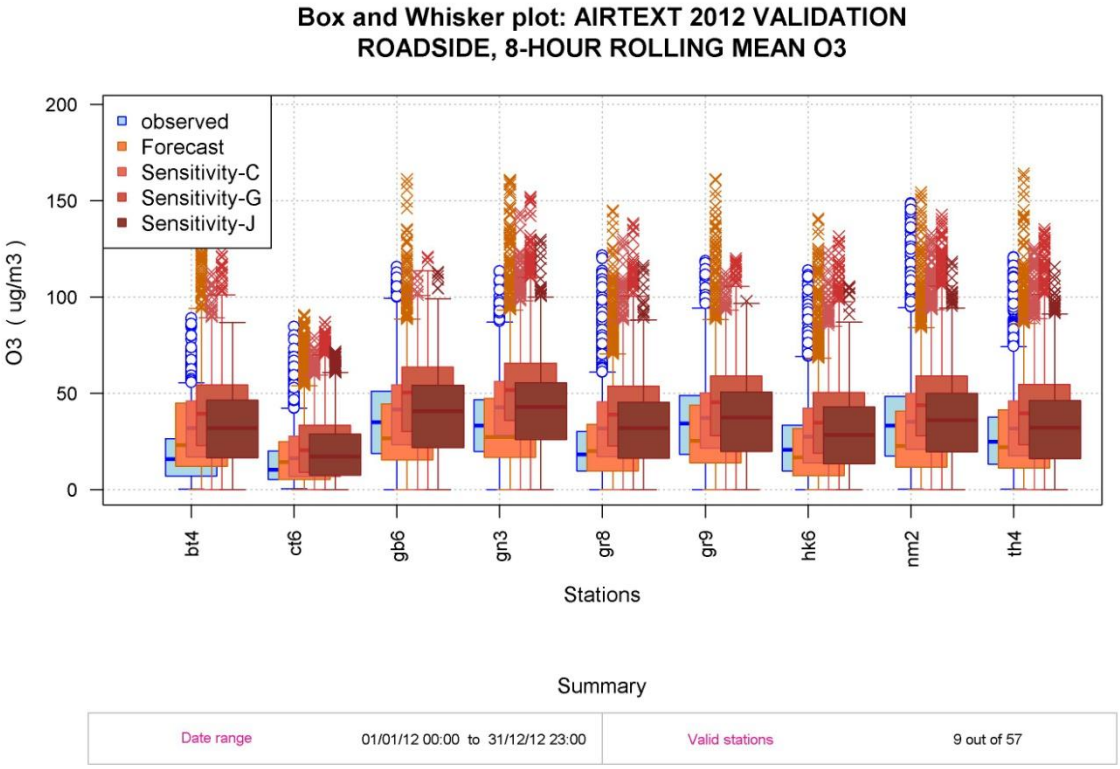


Figure 4.4 Example of a box and whisker plot from the concentration evaluation part of the Model Evaluation tool (grouped by station and filtered by model and pollutant; displaying single station type 'Roadside' for pollutant O_3)

4.2.1.4 Scatter Plot

This plot compares the modelled and observed concentrations on a scatter plot. The frequency scatter plot shows the frequency of occurrence of each point whereas the conventional scatter plot shows one point per pairwise modelled-observed data point. The data can optionally be plotted on a $\log_{10}$ scale. The black solid line is the 1:1 line. The dotted lines are the factor of 2 lines. Figure 4.5 presents an example frequency scatter plot and Figure 4.6 presents an example conventional scatter plot.

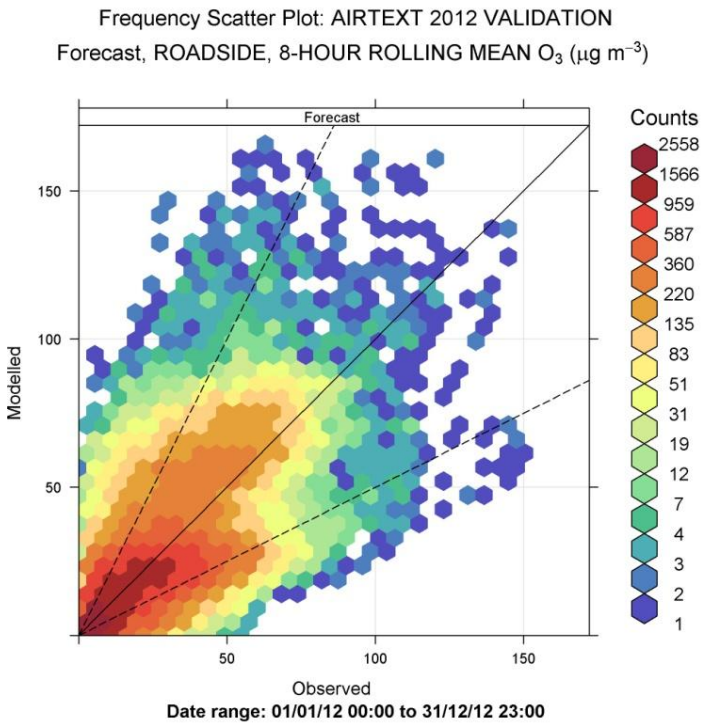


Figure 4.5 Example of a frequency scatter plot from the concentration evaluation part of the Model Evaluation tool (grouped by station and filtered by model and pollutant; displaying single model 'Forecast' and single station type 'Roadside' for pollutant O_3)

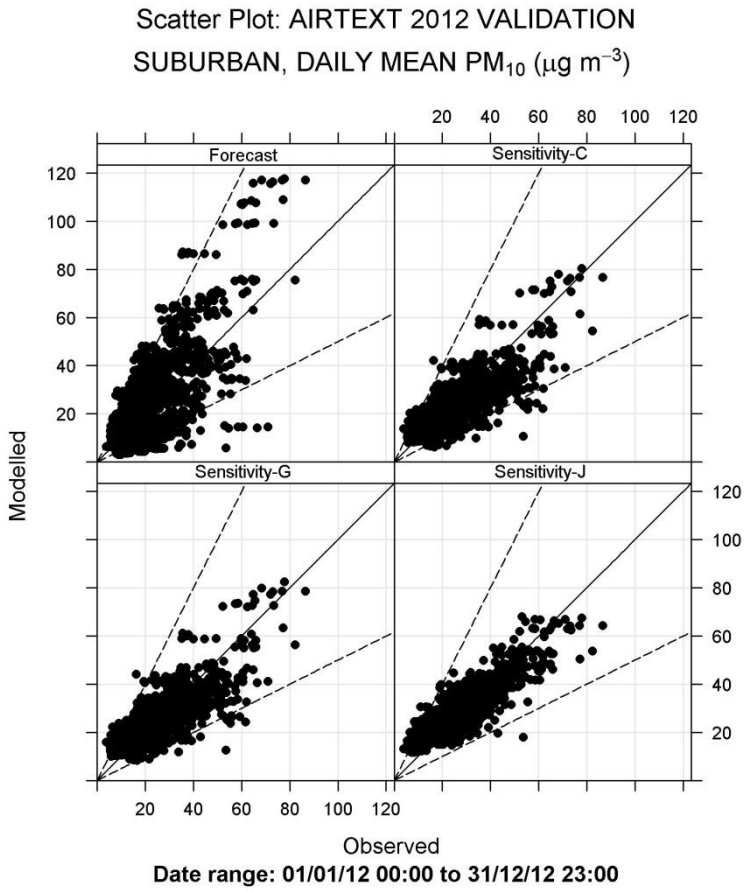


Figure 4.6 Example of a conventional scatter plot from the concentration evaluation part of the Model Evaluation tool (grouped by station and filtered by model and pollutant; displaying single station type 'Industrial' and pollutant PM_{10})

4.2.1.5 *Quantile-Quantile Plot*

This plot compares the modelled and observed concentrations ordered independently from lowest to highest concentration, as a quantile-quantile plot. The data can optionally be plotted on $\log_{10}$ scales. The black solid line is the 1:1 line. The dotted lines are the factor of 2 lines. Figure 4.7 presents an example quantile-quantile plot.

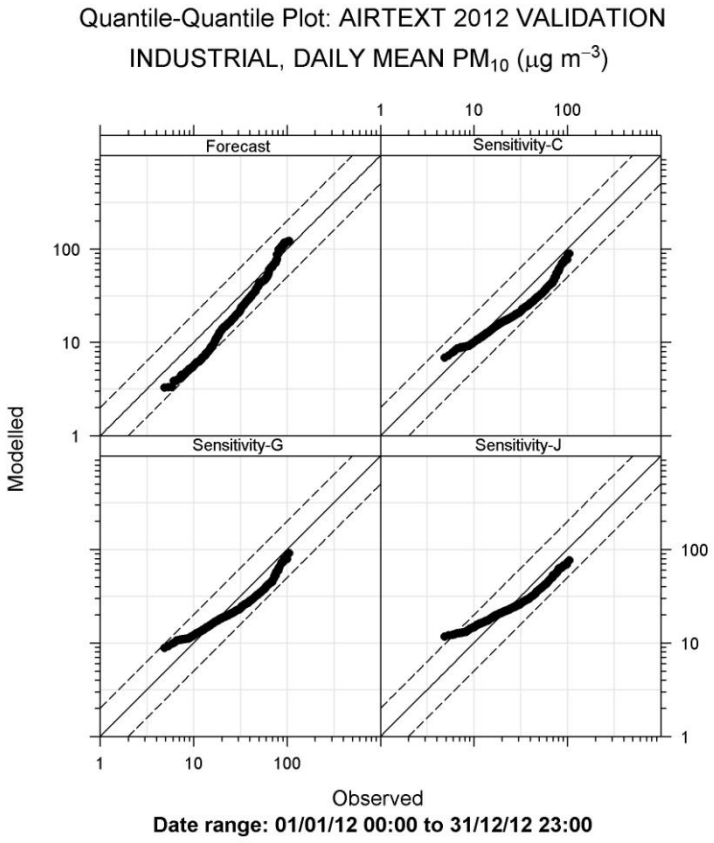


Figure 4.7 Example of a quantile-quantile plot from the concentration evaluation part of the Model Evaluation tool (grouped by station and filtered by model and pollutant; displaying pollutant PM_{10} on a $\log_{10}$ scale)

4.2.1.6 *CSV output file*

When the "output processed data and statistics as CSV" option is checked, statistics are calculated for each pollutant (in output units) and are output in a CSV file, *<*_conc_stats.csv>*. The statistics are calculated for the variable by which the data is grouped and additionally over all data, filtered by the filters selected. For example, if data is grouped by station and filtered by model and pollutant, the statistics are output for each station and for all valid stations in each model and pollutant.

The statistics that are output in are given in Table 4.1. Where 'obs' or 'mod' occur in variable names, these indicate observed or modelled values respectively.

Name	Description	Equation
Num.valid.values	Number of values	
obs.mean mod.mean	Mean	$1/n \sum C$
SDO SDM	Standard Deviation	$\sqrt{1/n \sum (C - \overline{C})^2}$
MB	Mean Bias	$\overline{(C_p - C_o)}$
NMSE	Normalised Mean-Square-Error	$\overline{(C_p - C_o)^2} / \overline{C_o C_p}$
R	Pearson's Correlation Coefficient	$\mathrm{cov}(C_p, C_o)/\sigma_{C_p}\sigma_{C_o}$
Fac2	Factor of 2	Fraction of data where $0.5 \leq C_p/C_o \leq 2$ (when $C_o = 0$, $C_p/C_o \to \infty$ and the data pair is not counted)
Fb	Fractional Bias	$(\overline{C_p} - \overline{C_o})/0.5(\overline{C_o} + \overline{C_p})$
Fs	Fractional Standard Deviation	$(\sigma_{C_p} - \sigma_{C_o})/0.5(\sigma_{C_o} + \sigma_{C_p})$
obs.max mod.max	Maximum	$\max C$
obs.RHC mod.RHC	Robust Highest Concentration	$\chi(n) + (\chi - \chi(n))\ln\left(\frac{3n-1}{2}\right)$, where n is the number of values used to characterise the upper end of the concentration distribution, χ is the average of the $n-1$ largest values, and $\chi(n)$ is the n^{th} largest value; n is taken to be 26.

Table 4.1 Details of the statistics output by the concentration evaluation part of the Model Evaluation tool

In addition, some statistics for each station that are presented on the target plots and the box plots are also output in the CSV file for each station:

1. Lower whisker (observed)
2. Lower quartile (observed)
3. Median (observed)
4. Upper quartile (observed)
5. Upper whisker (observed)
6. Lower whisker (modelled)

7. Lower quartile (modelled)
8. Median (modelled)
9. Upper quartile (modelled)
10. Upper whisker (modelled)
11. centralised root-mean-square-error (CRMSE)
12. root-mean-square measurement uncertainty (RMSu)
13. normalised mean standard deviation (NMSD)
14. CRMSE.sign.DELTA.1.2 (CRMSE multiplied by +1 if SDO>SDM and multiplied by -1 if SDM>SDO)
15. Target parameter (T.DELTA.1.2)
16. CRMSE.sign.DELTA.3.3. (CRMSE multiplied by +1 if $\mathrm{NMSD} > \sqrt{2(1-R)}$ and multiplied by -1 if $\mathrm{NMSD} < \sqrt{2(1-R)}$)
17. Target parameter (T.DELTA.3.3)

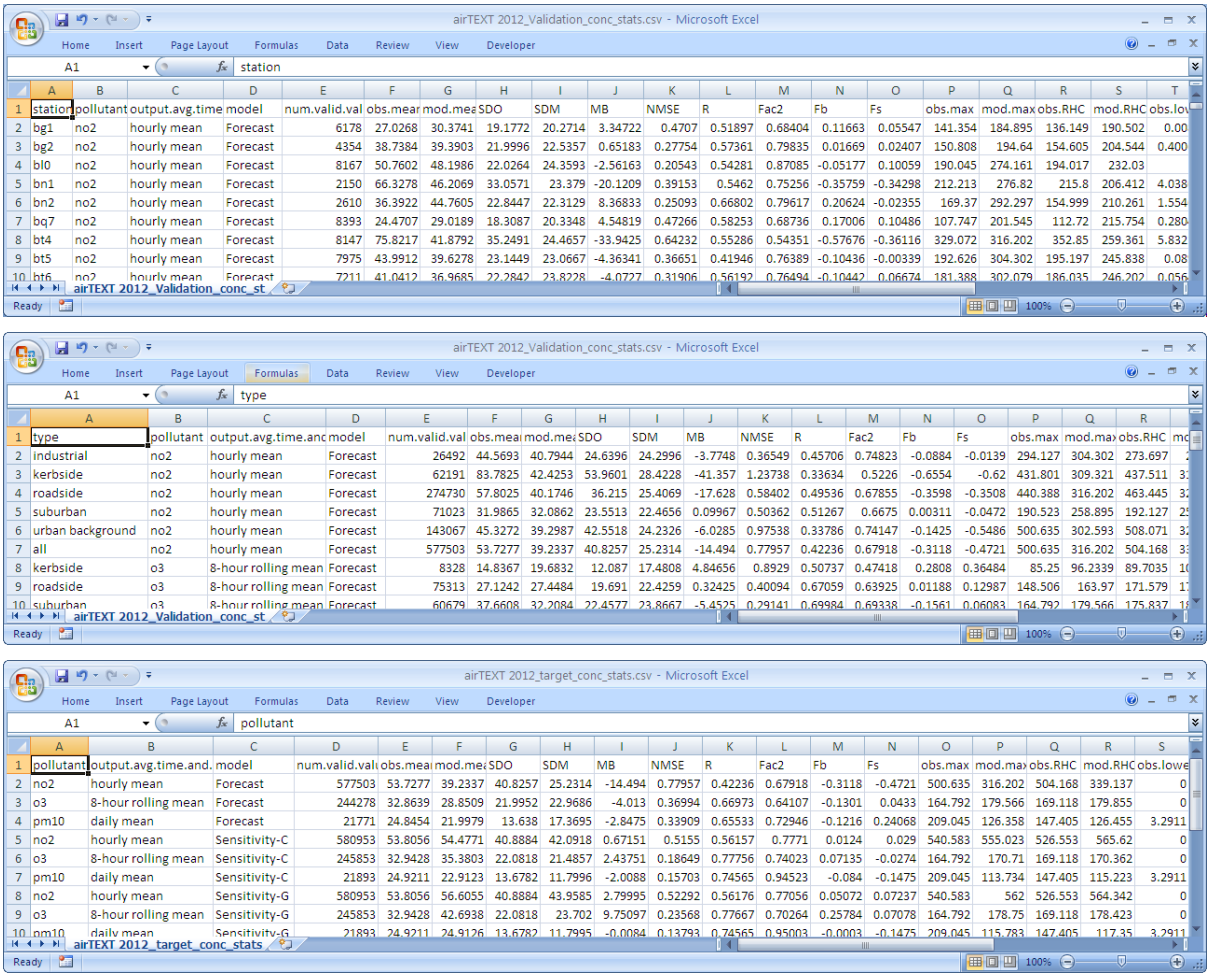


Figure 4.8 Examples of concentration evaluation output CSV file (grouped by station, station type and pollutant and filtered by model and pollutant)

4.2.2 Forecast Index Evaluation Output

The forecast index evaluation part of the model evaluation tool produces three types of output; these are described in this section.

All three types of graph are based on a 'forecast index', which is calculated from the modelled and observed data according to the index threshold definitions in the index scales file (described in section 3.3, step 8).

4.2.2.1 Forecast index accuracy

This part of the tool assesses the performance of the model's forecast index predictions against forecast indices calculated from observed concentrations.

The graph is a stacked bar chart that shows, for each station, the percentage of calculated forecast indices valid for comparison where the modelled index was equal to the observed index (green) and where the modelled index was equal to the observed index plus or minus one band (grey). Only forecast periods for which both modelled and observed indices can be calculated are included in the assessment. The stations are sorted by the number of indices valid for comparison, which is also shown on the chart by the blue circles and the right-hand y-axis. A key to the stations is given below the graph. Refer to Figure 4.9 for an example.

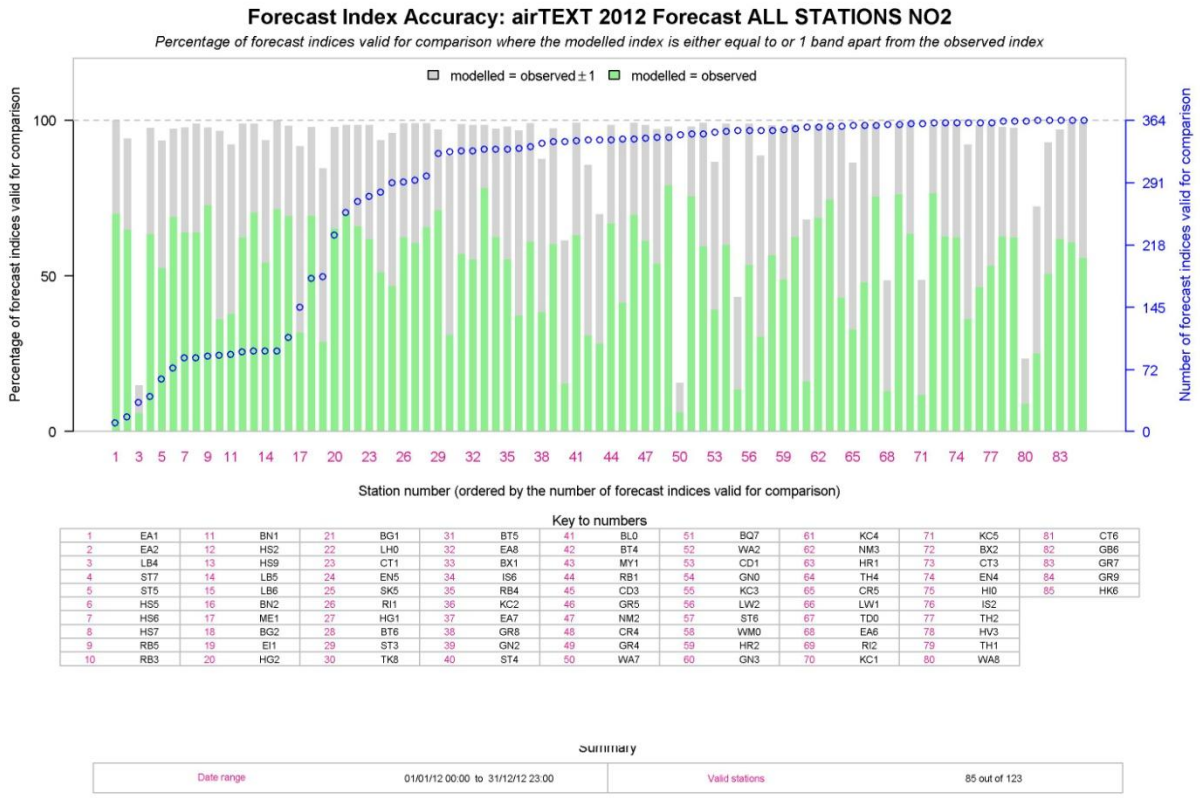


Figure 4.9 Example of the forecast index accuracy graph (grouped by station and filtered by model and pollutant; displaying model 'Forecast' and pollutant NO_2)

4.2.2.2 *Forecast alert accuracy*

It is usual in operational pollution forecasting to use pollution bandings to help communicate pollution levels to the public. For example, a common set of bandings for a 1 to 10 forecast index scale is shown in Table 4.2.

Band	Forecast index range	Alert threshold
LOW	1 to 3	n/a
MODERATE	4 to 6	4
HIGH	7 to 9	7
VERY HIGH	10	10

Table 4.2 Example set of bandings, with associated alert thresholds

Depending on the system, a forecast index in the MODERATE, HIGH or VERY HIGH range may trigger an alert to the public; it is therefore important for system operators to understand whether the system issues these alerts correctly.

The assessment of forecast alerts is carried out by calculating metrics for each monitoring station based on considering the exceedence of an alert threshold as an 'event'. The number of events observed and modelled, modelled but not observed, observed but not modelled, not modelled and not observed are summed to get the parameters *a*, *b*, *c* and *d* respectively. This is summarised in Table 4.3.

		Event observed	
		Yes	No
Event modelled	Yes	*a*	*b*
	No	*c*	*d*

Table 4.3 Definition of the forecast alert parameters

The forecast index evaluation includes two sets of graphical output for the assessment of the accuracy of forecast alerts:

1. **Odds ratio skill score (ORSS)**

   The odds ratio skill score (ORSS) is calculated from the alert metrics as follows:

$$\text{Odds ratio (OR)} = \frac{ad}{bc}$$

$$\text{ORSS} = \frac{OR - 1}{OR + 1}$$

   A perfect system will have *b* and *c* equal to zero, which means $\text{OR} \to \infty$, which means $\text{ORSS} \to 1$.

   A poor system will have *a* and *d* equal to zero, which means $\text{OR} = 0$, which means $\text{ORSS} = -1$.

   The odds ratio is a good metric for determining if a model is good at correctly issuing and not issuing alerts. It gives equal weighting to the correct prediction of an alert and to the correct non-prediction of a non-alert. If no alerts are observed or no alerts are forecast then ORSS is invalid. The graph shows ORSS for each station, where the stations are ordered by the number of observed alert threshold exceedences. The number of observed and modelled alerts for each station is also plotted in blue (right-hand y-axis). A key to the stations is shown below the graph. Figure 4.10 shows an example of an ORSS graph.

Figure 4.10 Example of a ORSS graph produced by the forecast index evaluation (grouped by pollutant and filtered by model; displaying model 'Forecast')

2. Performance metrics

In an operational pollution forecasting system, it is important not to issue alerts when there should not be an alert, but it is arguably *more* important to accurately issue an alert when an alert should be issued.

The following performance metrics give information about the skill of a model in terms of its ability to issue accurate alerts, and any tendency towards 'false alarms'.

$$\text{Probability of a correct forecast (PCF)} = \frac{a+d}{a+b+c+d}$$

$$\text{Probability of detection (POD)} = \frac{a}{a+c}$$

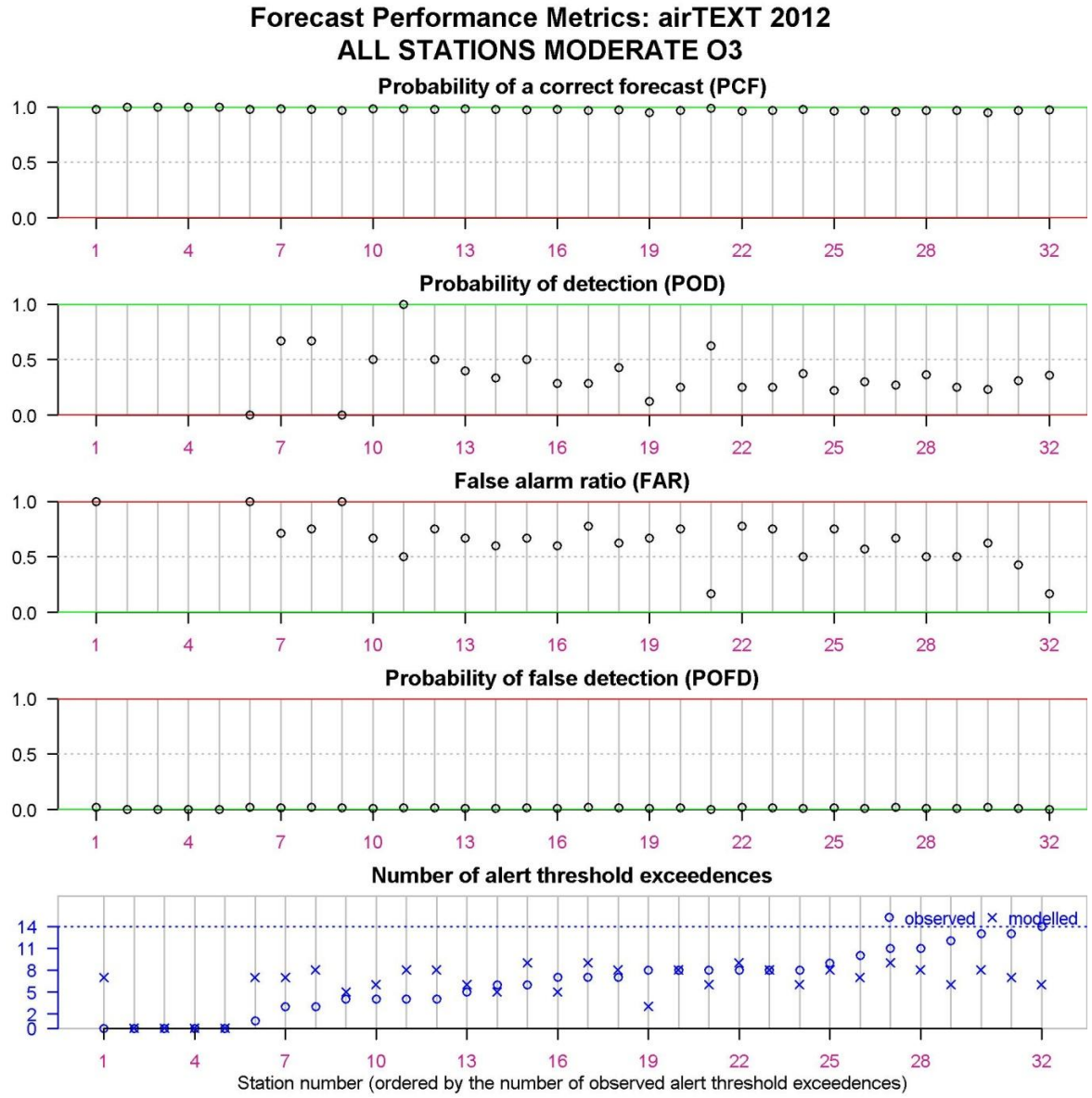
$$\text{False alarm ratio (FAR)} = \frac{b}{a+b}$$

$$\text{Probability of false detection (POFD)} = \frac{b}{b+d}$$

Each of these metrics lies in the range 0 to 1. A good score for PCF and POD is 1; a good score for FAR and POFD is 0.

The graphical output from the performance metrics option is a series of five graphs; the four metrics described above and a graph showing the number of observed and modelled alert threshold exceedences for each station. Again, the stations are sorted by the number

of observed alert threshold exceedences and a key to the stations is shown below the graphs. Figure 4.11 and Figure 4.12 show examples of this output grouped and filtered by different variables.



Key to numbers

1	BT4	11	LH0	21	HI0	31	TH1
2	CT6	12	TH4	22	RB1	32	RI2
3	EA1	13	GR9	23	ST3		
4	HS2	14	EA7	24	WM0		
5	MY1	15	TD0	25	GR4		
6	GN3	16	BL0	26	NM3		
7	HK6	17	BX1	27	BQ7		
8	LW1	18	WA2	28	NM2		
9	CR3	19	CT1	29	KC1		
10	GR8	20	GB6	30	HG2		

Summary

Date range	01/01/12 00:00 to 31/12/12 23:00	Valid stations	32 out of 123

Figure 4.11 Example of a forecast performance metrics graph produced by the forecast index evaluation (grouped by station and filtered by model and pollutant; displaying pollutant O3)

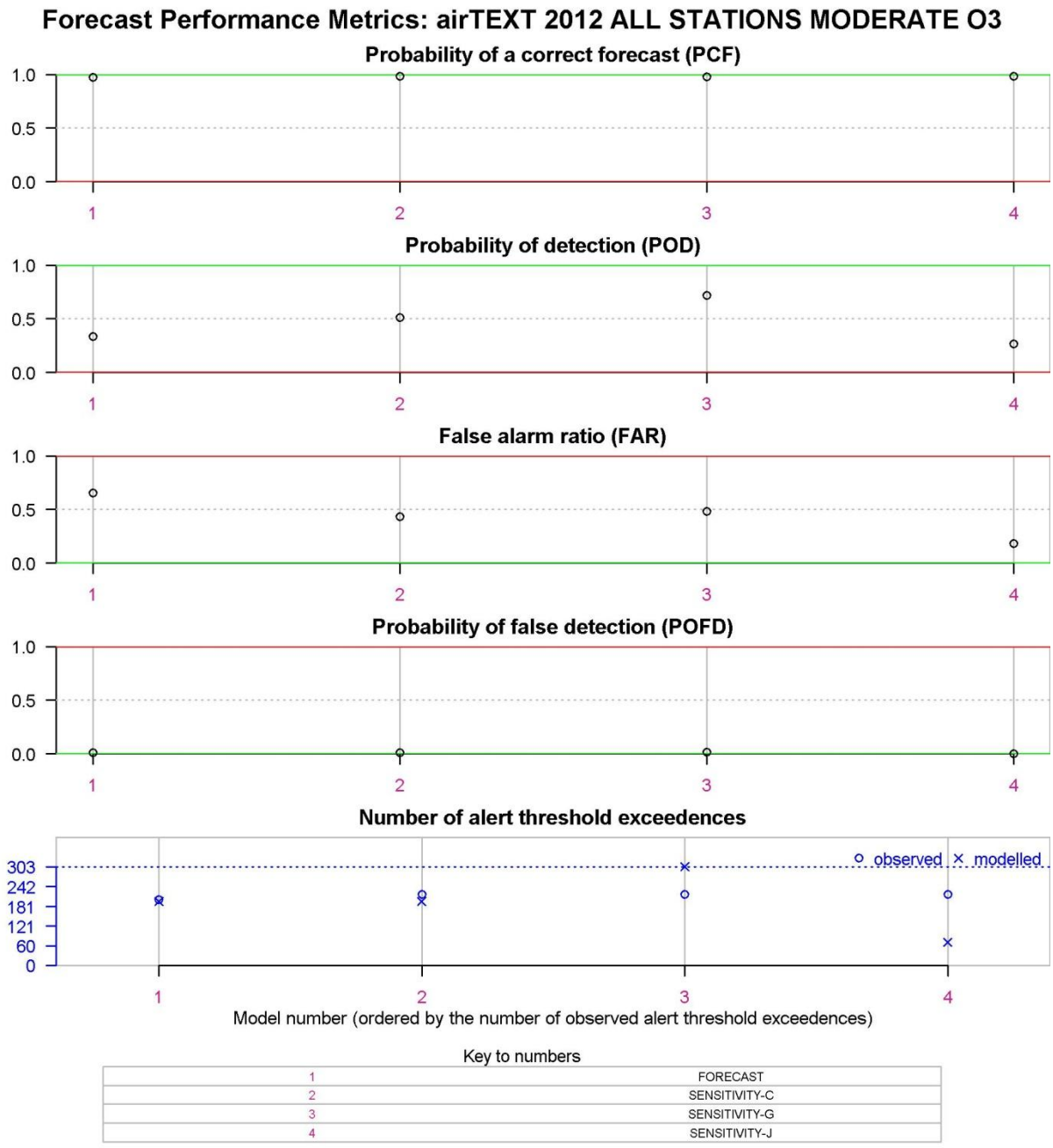
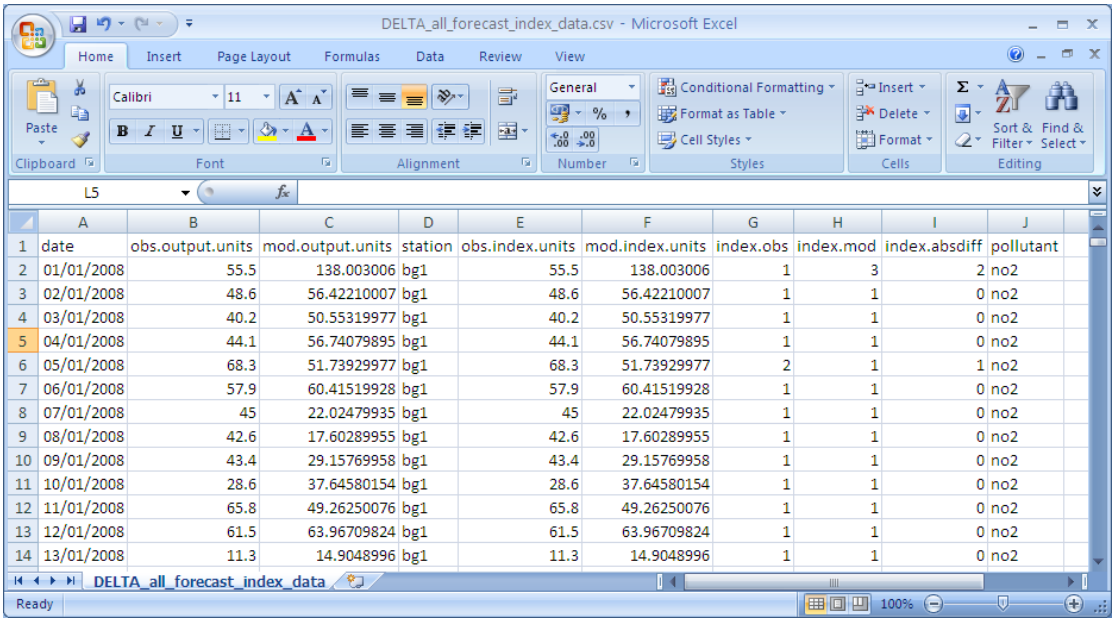


Figure 4.12 Example of a forecast performance metrics graph produced by the forecast index evaluation (grouped by model and filtered by pollutant; displaying pollutant O3)

4.2.2.3 CSV output files

*_forecast_index_data.csv:

This contains the modelled and observed concentrations using the index averaging time and statistic, in both output units and index units, also the observed and modelled index value for each averaging period and the absolute difference between the observed and modelled indices. For an example refer to Table 4.4.

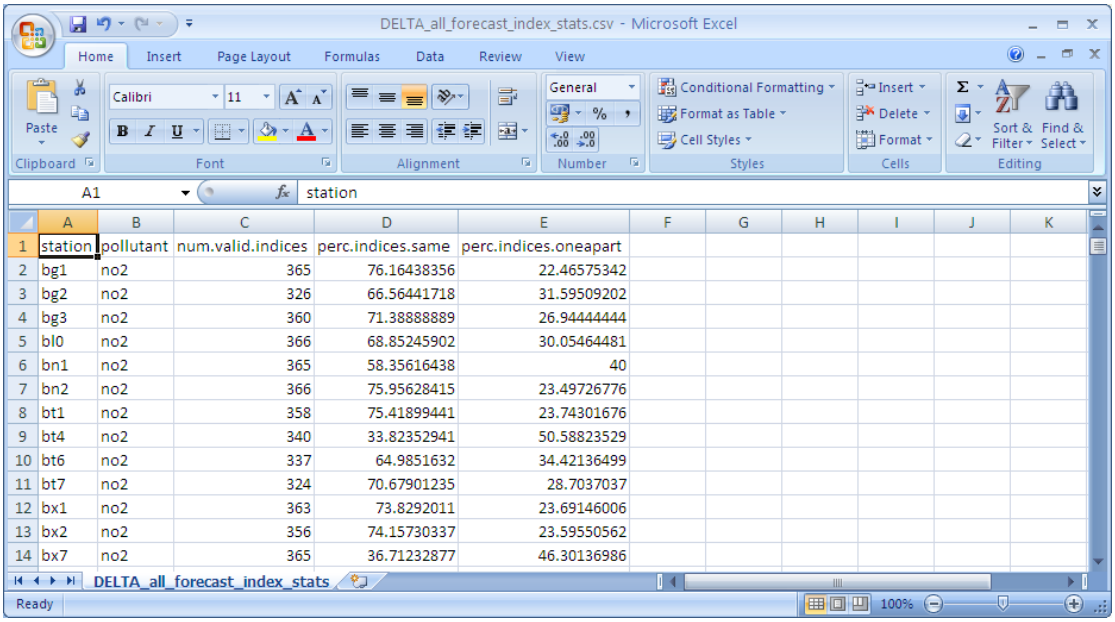


date	obs.output.units	mod.output.units	station	obs.index.units	mod.index.units	index.obs	index.mod	index.absdiff	pollutant
01/01/2008	55.5	138.003006	bg1	55.5	138.003006	1	3	2	no2
02/01/2008	48.6	56.42210007	bg1	48.6	56.42210007	1	1	0	no2
03/01/2008	40.2	50.55319977	bg1	40.2	50.55319977	1	1	0	no2
04/01/2008	44.1	56.74079895	bg1	44.1	56.74079895	1	1	0	no2
05/01/2008	68.3	51.73929977	bg1	68.3	51.73929977	2	1	1	no2
06/01/2008	57.9	60.41519928	bg1	57.9	60.41519928	1	1	0	no2
07/01/2008	45	22.02479935	bg1	45	22.02479935	1	1	0	no2
08/01/2008	42.6	17.60289955	bg1	42.6	17.60289955	1	1	0	no2
09/01/2008	43.4	29.15769958	bg1	43.4	29.15769958	1	1	0	no2
10/01/2008	28.6	37.64580154	bg1	28.6	37.64580154	1	1	0	no2
11/01/2008	65.8	49.26250076	bg1	65.8	49.26250076	1	1	0	no2
12/01/2008	61.5	63.96709824	bg1	61.5	63.96709824	1	1	0	no2
13/01/2008	11.3	14.9048996	bg1	11.3	14.9048996	1	1	0	no2

Table 4.4 Example of forecast index evaluation output *_forecast_index_data.csv file

*_forecast_index_stats.csv:

This contains, for each pollutant and station, the number of valid calculated index values, the percentage of indices where the modelled index was correct and the percentage of indices where the modelled index was one band either above or below the correct index. For an example refer to Table 4.5.

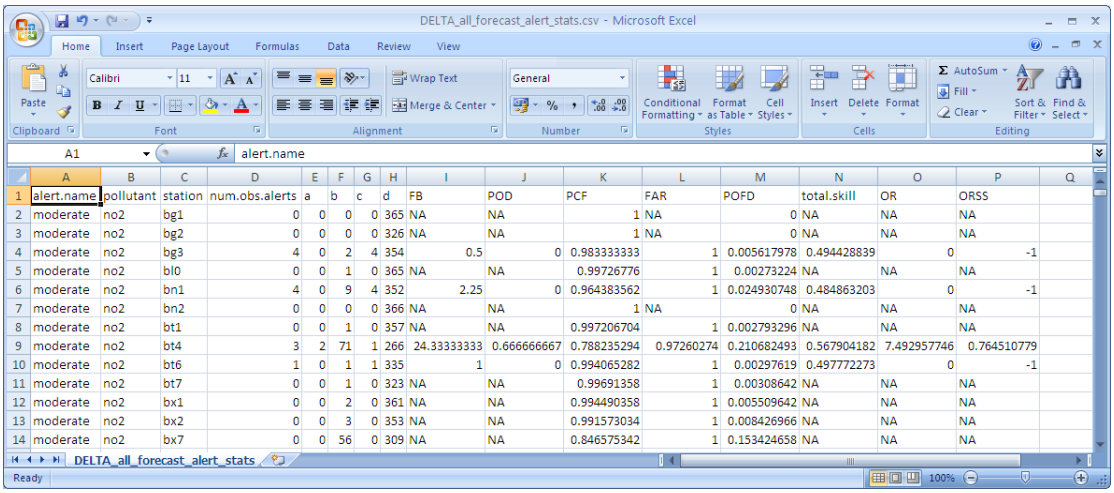


station	pollutant	num.valid.indices	perc.indices.same	perc.indices.oneapart
bg1	no2	365	76.16438356	22.46575342
bg2	no2	326	66.56441718	31.59509202
bg3	no2	360	71.38888889	26.94444444
bl0	no2	366	68.85245902	30.05464481
bn1	no2	365	58.35616438	40
bn2	no2	366	75.95628415	23.49726776
bt1	no2	358	75.41899441	23.74301676
bt4	no2	340	33.82352941	50.58823529
bt6	no2	337	64.9851632	34.42136499
bt7	no2	324	70.67901235	28.7037037
bx1	no2	363	73.8292011	23.69146006
bx2	no2	356	74.15730337	23.59550562
bx7	no2	365	36.71232877	46.30136986

Table 4.5 Example of forecast index evaluation output *_forecast_index_stats.csv file

*_forecast_alert_stats.csv:

This contains, for each station, for each pollutant, and for each alert threshold, the number of observed alerts, the 4 event parameters a, b, c and d, the performance metrics PCF, POD, FAR, POFD and also the odds ratio OR and the odds ratio skill score ORSS. Refer to Table 4.6 for an example.



alert.name	pollutant	station	num.obs.alerts	a	b	c	d	FB	POD	PCF	FAR	POFD	total.skill	OR	ORSS
moderate	no2	bg1	0	0	0	0	365	NA	NA	1	NA	0	NA	NA	NA
moderate	no2	bg2	0	0	0	0	326	NA	NA	1	NA	0	NA	NA	NA
moderate	no2	bg3	4	0	2	4	354	0.5	0	0.983333333	1	0.005617978	0.494428839	0	-1
moderate	no2	bl0	0	0	1	0	365	NA	NA	0.99726776	1	0.00273224	NA	NA	NA
moderate	no2	bn1	4	0	9	4	352	2.25	0	0.964383562	1	0.024930748	0.484863203	0	-1
moderate	no2	bn2	0	0	0	0	366	NA	NA	1	NA	0	NA	NA	NA
moderate	no2	bt1	0	0	1	0	357	NA	NA	0.997206704	1	0.002793296	NA	NA	NA
moderate	no2	bt4	3	2	71	1	266	24.33333333	0.666666667	0.788235294	0.97260274	0.210682493	0.567904182	7.492957746	0.764510779
moderate	no2	bt6	1	0	1	1	335	1	0	0.994065282	1	0.00297619	0.497772273	0	-1
moderate	no2	bt7	0	0	1	0	323	NA	NA	0.99691358	1	0.00308642	NA	NA	NA
moderate	no2	bx1	0	0	2	0	361	NA	NA	0.994490358	1	0.005509642	NA	NA	NA
moderate	no2	bx2	0	0	3	0	353	NA	NA	0.991573034	1	0.008426966	NA	NA	NA
moderate	no2	bx7	0	0	56	0	309	NA	NA	0.846575342	1	0.153424658	NA	NA	NA

Table 4.6 Example of forecast index evaluation output *_forecast_alert_stats.csv file

4.3 Model Diagnostics Tool Output

4.3.1 Time Variation Plot

This compares modelled and observed concentrations by (clockwise from top) hour and day of the week, day of the week, month of the year and hour of the day. The shaded area indicates the 95% confidence interval in the mean. For an example, refer to Figure 4.13.

4.3.2 Scatter Plot

This compares the modelled and observed concentrations on a scatter plot, optionally filtered by weekday, month or season. The frequency scatter plot shows the frequency of occurrence of each point whereas the conventional scatter plot shows one point per pairwise modelled-observed data point. When the calculated forecast indices are chosen in the data to plot, the output defaults to a frequency scatter plot. The black solid line is the 1:1 line. The dotted lines are the factor of 2 lines. Figures Figure 4.14 to Figure 4.17 show examples of frequency and conventional scatter plots for different 'data to plot' and different filtering options.

4.3.3 Time Plot

This plots a time series of modelled and observed concentrations. Figures Figure 4.18 and Figure 4.19 show examples of a time plot for different 'data to plot'.

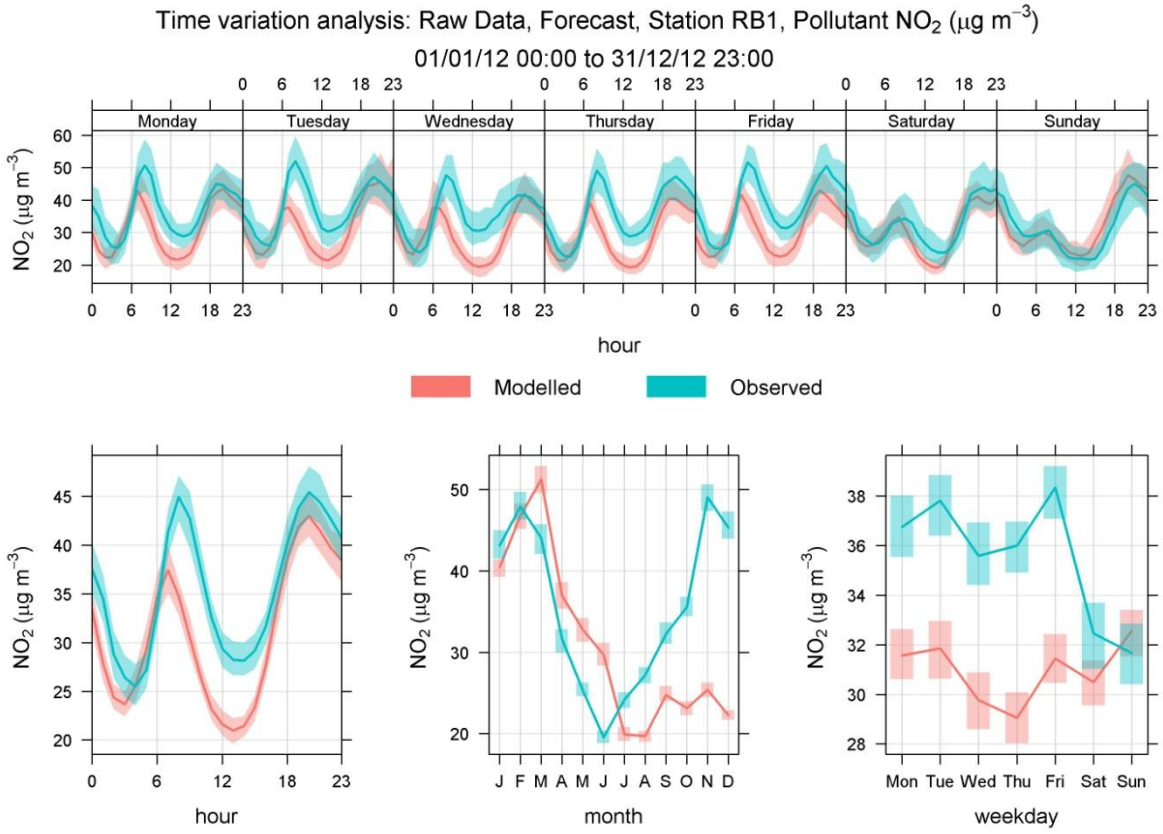


Figure 4.13 Example of a time variation plot from the Model Diagnostics tool.

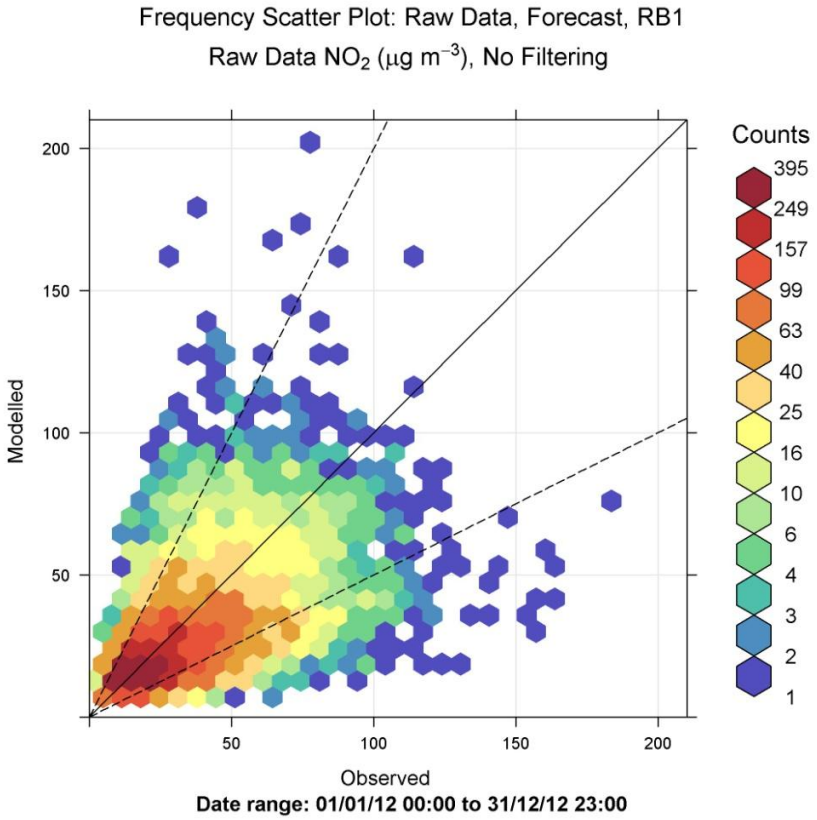


Figure 4.14 Example of a frequency scatter plot from the Model Diagnostics tool; raw data with no filtering

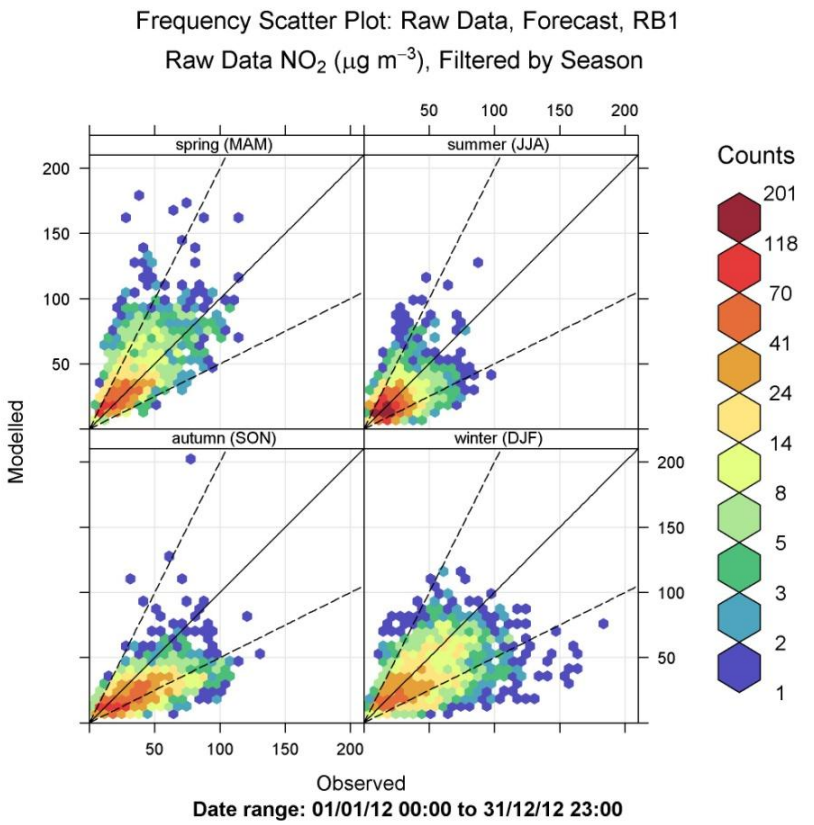


Figure 4.15 Example of a frequency scatter plot from the Model Diagnostics tool; raw data filtered by season

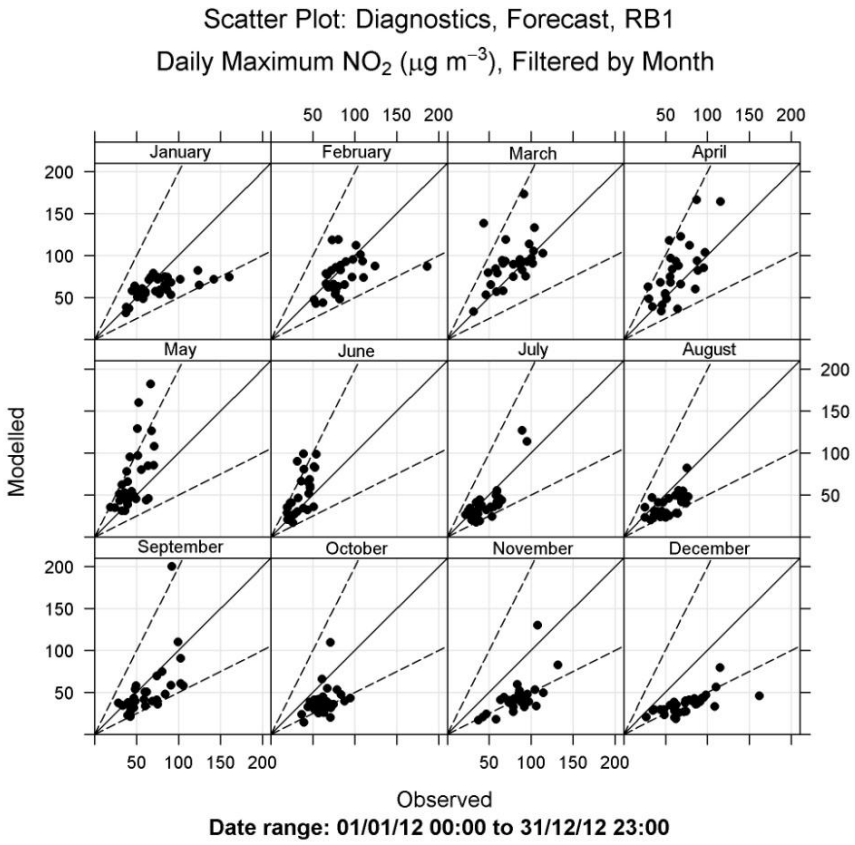


Figure 4.16 Example of a conventional scatter plot from the Model Diagnostics tool; calculated statistics (daily maximum NO_2) filtered by month

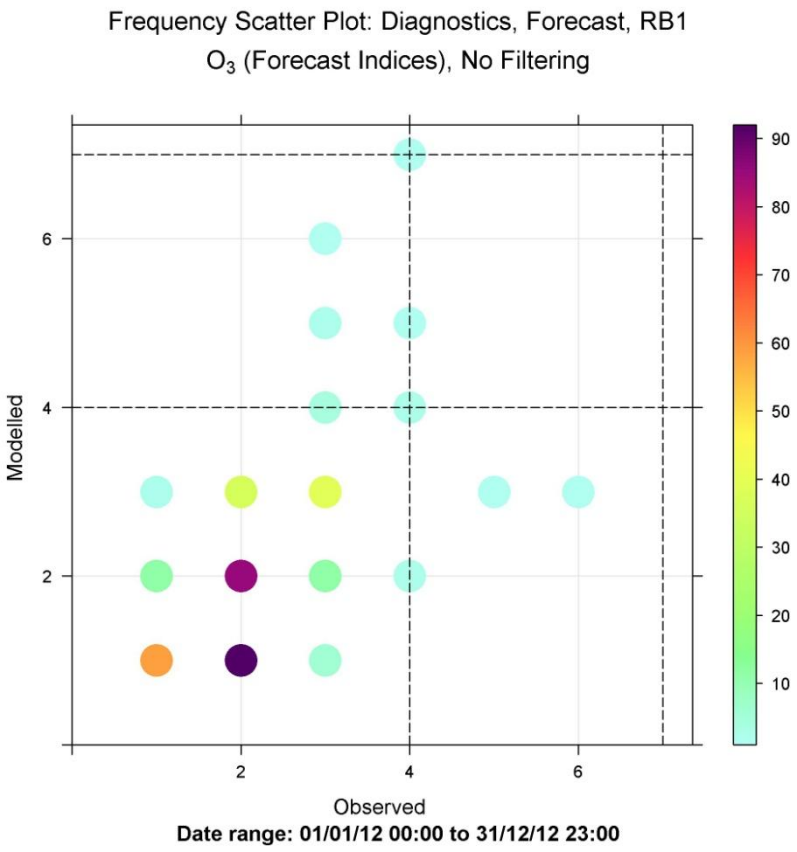


Figure 4.17 Example of a frequency scatter plot from the Model Diagnostics tool; forecast indices with no filtering

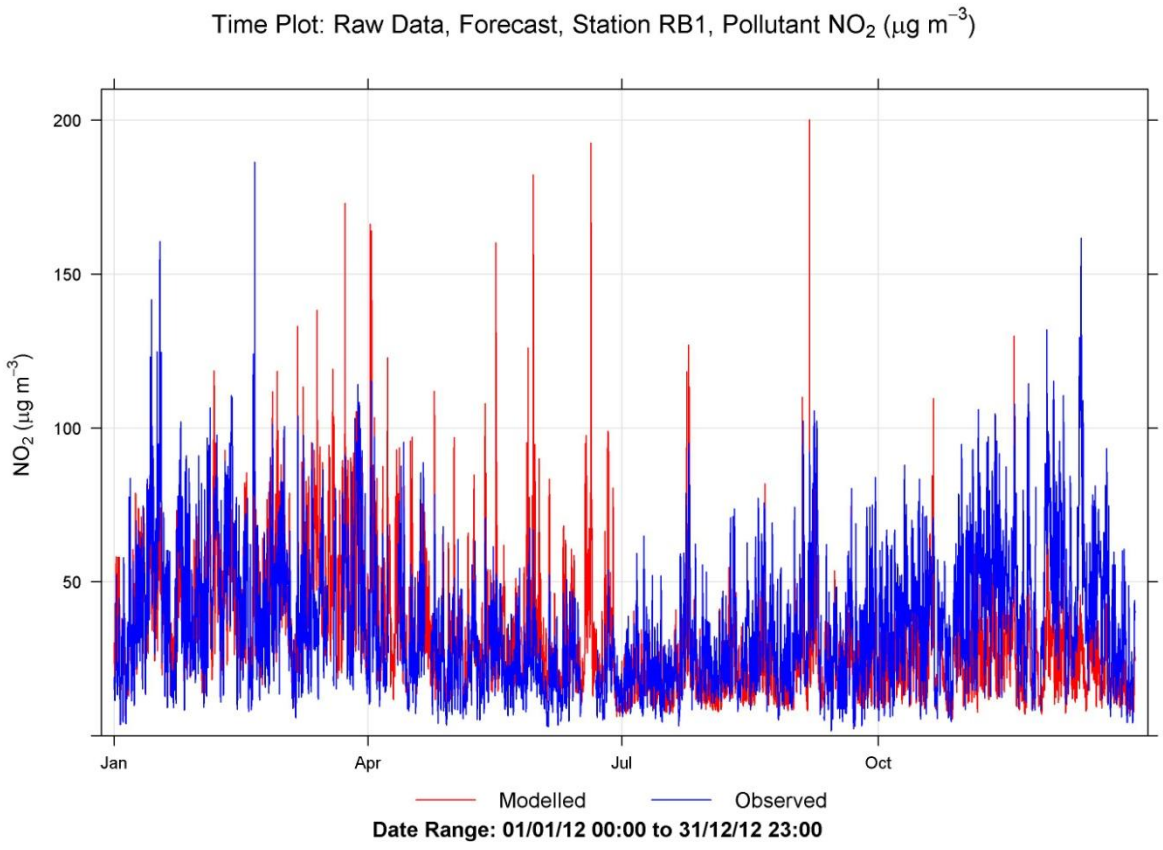


Figure 4.18 Example of a time plot from the Model Diagnostics tool; raw data

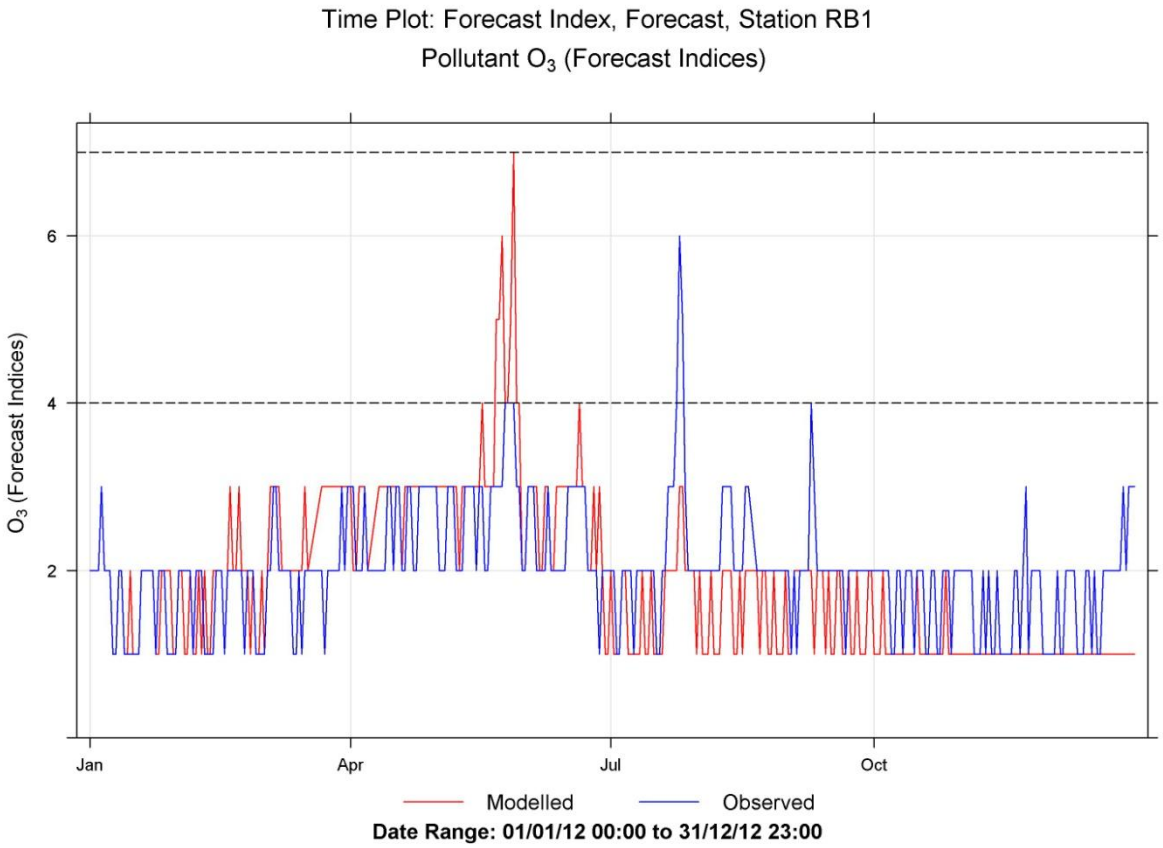
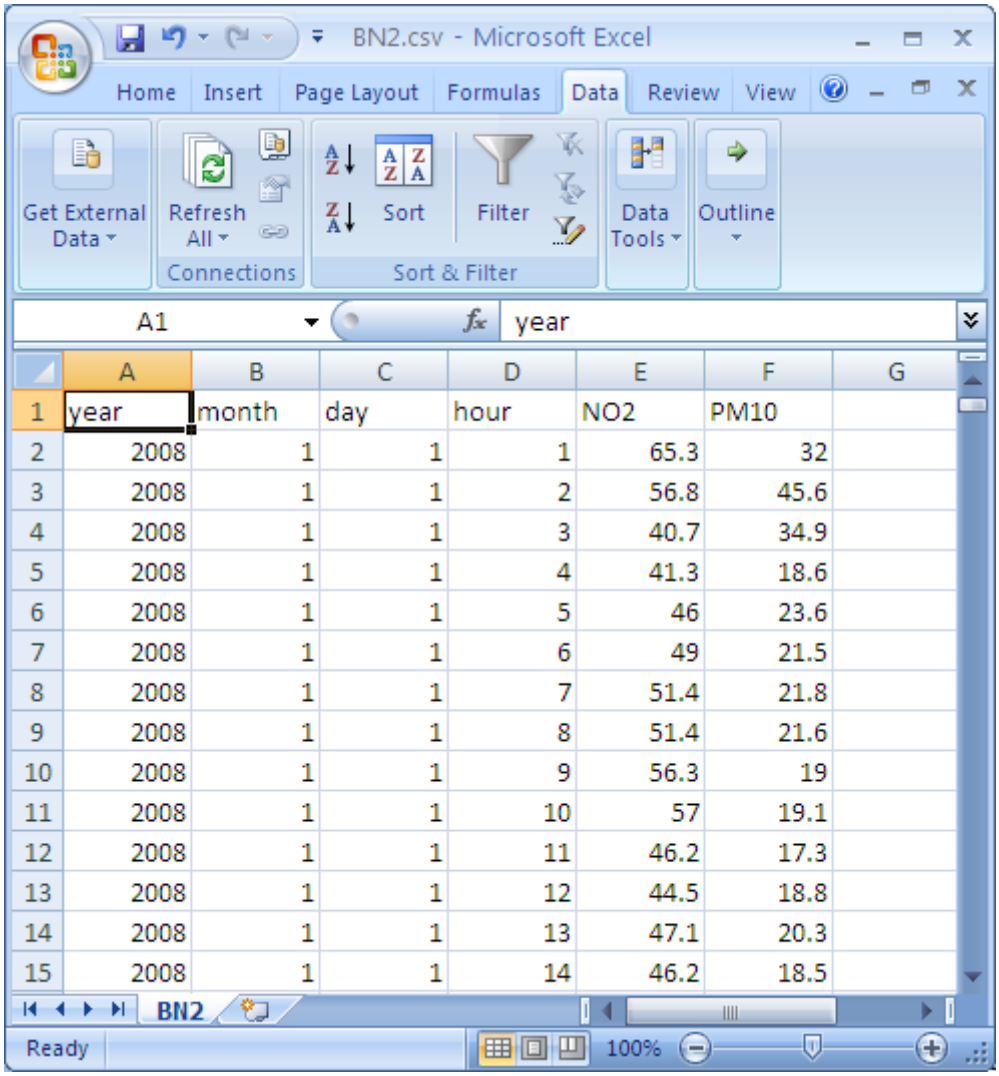


Figure 4.19 Example of a time plot from the Model Diagnostics tool; forecast indices

5 File formats

5.1 CSV modelled and observed data



	A	B	C	D	E	F	G
1	year	month	day	hour	NO2	PM10	
2	2008	1	1	1	65.3	32	
3	2008	1	1	2	56.8	45.6	
4	2008	1	1	3	40.7	34.9	
5	2008	1	1	4	41.3	18.6	
6	2008	1	1	5	46	23.6	
7	2008	1	1	6	49	21.5	
8	2008	1	1	7	51.4	21.8	
9	2008	1	1	8	51.4	21.6	
10	2008	1	1	9	56.3	19	
11	2008	1	1	10	57	19.1	
12	2008	1	1	11	46.2	17.3	
13	2008	1	1	12	44.5	18.8	
14	2008	1	1	13	47.1	20.3	
15	2008	1	1	14	46.2	18.5	

Figure 5.1 Example of an observed data set in CSV file format

Restrictions:

- Include no more than one column per pollutant to be analysed
- The date must be given by the year, month, day and hour.
- The station column is optional. If it is present, the header should read “station”; if it is missing (as in the screenshot above), then all the data in that file is assigned to a ‘station’ with the name of the filename, e.g. BG1.csv would be assigned to station ‘BG1’.
- Modelled data must contain data for each pollutant to be analysed. Note that if multiple files are used, that means each pollutant to be analysed must be present in at least one file. Observed data does not have to include data for all required pollutants.

6 Batch mode facility

The Toolkit is supplied with R-script versions of the Data Input, Model Evaluation and Model Diagnostics tools which can be run from the command-line, making it possible to automate tasks, for example in a batch file.

These R-script files can be found in the 'Rscripts' sub-directory of the Toolkit installation directory.

The basic procedure is the same for each of the above tools:

1. Run the tool as usual from the interface

   As well as generating all the usual output this saves a .ini file containing all the information and settings that you entered in the interface. This .ini file is saved in the output directory with the specified output label.

2. To re-run the tool with the same information and settings from the command line, type the following on the command line or in a batch file:

   {Rscript exe path} {R-script path} {ini file path}

   The text output that usually appears in the output screen of the interface is saved instead in an .out file with the same name as the .ini file.

{Rscript exe path}: This is the pathname of the Rscript.exe file installed with R. This must be used when running R from the command line. A typical path on a Windows PC (where R was installed in "C:\Program Files\R\R-2.15.3") would be

"C:\Program Files\R\R-2.15.3\bin\i386\Rscript.exe"

{Rscript path}: This is the pathname of the R-script for the tool you want to run. For example, for the Data Input tool, if the Toolkit is installed at "C:\Myair Toolkit v3.0" then the path would be

"C:\Myair Toolkit v3.0\RScripts\DSLocal_DataInput.r"

{ini file path}: This is the path of the .ini file that was saved when you ran the tool from the interface. For example, if you chose the output directory "C:\Toolkit_output" and the output label 'DELTA' then the .ini file path would be

"C:\Toolkit_output\DELTA.ini"

7 References

1. Thunis P., E. Georgieva and S. Galmarini, 2011. **A procedure for air quality models benchmarking**. http://fairmode.ew.eea.europa.eu/models-benchmarking-sg4/wg2_sg4_benchmarking_v2.pdf

2. David Carslaw and Karl Ropkins (2013). **openair: Open-source tools for the analysis of air pollution data**. R package version 0.8-5. http://www.openair-project.org/

3. Thunis, P., A. Pederzoli, E. Giorgieva, C. Cuvelier and D. Pernigotti (2013). **The DELTA tool and Benchmarking Report template: Concepts and User guide version 3.2.** http://aqm.jrc.ec.europa.eu/DELTA/data/DELTA_UserGuide_V3.pdf

4. Thunis, P., D. Pernigotti and M. Gerboles (2013). **Model quality objectives based on measurement uncertainty. Part I: Ozone.** Atmospheric Environment, In Press. http://dx.doi.org/10.1016/j.atmosenv.2013.05.018

5. Thunis, P., D. Pernigotti, M. Gerboles and C. Belis (2013). **Model quality objectives based on measurement uncertainty. Part II: PM10 and NO2.** Atmospheric Environment, Submitted.

6. R Development Core Team (2010). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

7. **Ilhami Visne, Erkan Dilaveroglu, Klemens Vierlinger, Martin Lauss, Ahmet Yildiz, Andreas Weinhaeusel, Christa Noehammer, Friedrich Leisch and Albert Kriegner (2009). RGG: A general GUI Framework for R scripts**. BMC Bioinformatics. Vol 10, p 1471-2105. http://www.biomedcentral.com/1471-2105/10/74

8. RGG project website: http://rgg.r-forge.r-project.org/index.html